The Complexity of Nash Equilibria in Limit-Average Games*

Michael Ummels[1] and Dominik Wojtczak[2,3]

[1] LSV, CNRS & ENS Cachan, France
E-Mail: ummels@lsv.ens-cachan.fr
[2] University of Liverpool, UK
E-Mail: d.k.wojtczak@liv.ac.uk
[3] Oxford University Computing Laboratory, UK

Abstract. *We study the computational complexity of Nash equilibria in concurrent games with limit-average objectives. In particular, we prove that the existence of a Nash equilibrium in randomised strategies is undecidable, while the existence of a Nash equilibrium in pure strategies is decidable, even if we put a constraint on the payoff of the equilibrium. Our undecidability result holds even for a restricted class of concurrent games, where nonzero rewards occur only on terminal states. Moreover, we show that the constrained existence problem is undecidable not only for concurrent games but for turn-based games with the same restriction on rewards. Finally, we prove that the constrained existence problem for Nash equilibria in (pure or randomised) stationary strategies is decidable and analyse its complexity.*

1 Introduction

Concurrent games provide a versatile model for the interaction of several components in a distributed system where the components perform actions in parallel [17]. Classically, such a system is modelled by a family of concurrent two-player games, one for each component, where one component tries to fulfil its specification against the coalition of all other components. In practice, this modelling is often too pessimistic because it ignores the specifications of the other components. We argue that a distributed system is more faithfully modelled by a multiplayer game where each player has her own objective, which is independent of the other players' objectives.

Another objection to the classical theory of verification and synthesis has been that specifications are *qualitative*: either the specification is fulfilled, or it is violated. Examples of such specifications include reachability properties,

* This work was supported by ESF RNP "Games for Design and Verification" (GAMES), the French project ANR-06-SETI-003 (DOTS) and EPSRC grant EP/G050112/1.

where a certain set of target states has to be reached, or safety properties, where a certain set of states has to be avoided. In practice, many specifications are of a *quantitative* nature, examples of which include minimising average power consumption or maximising average throughput. Specifications of the latter kind can be expressed by assigning (positive or negative) rewards to states or transitions and considering the *limit-average* reward gained from an infinite play. In fact, concurrent games where a player's payoff is defined in such a way have been a central topic in game theory (see the related work section below).

The most common solution concept for games with multiple players is that of a Nash equilibrium [20]. In a Nash equilibrium, no player can improve her payoff by changing her strategy unilaterally. Unfortunately, Nash equilibria do not always exist in concurrent games, and if they exist, they may not be unique. In applications, one might look for an equilibrium where some players receive a high payoff while other players receive a low payoff. Formulated as a decision problem, given a game with k players and thresholds $\bar{x}, \bar{y} \in (\mathbb{Q} \cup \{\pm\infty\})^k$, we want to know whether the game has a Nash equilibrium whose payoff lies in-between $\bar{x}$ and $\bar{y}$; we call this decision problem NE.

The problem NE comes in several variants, depending on the type of strategies one considers: On the one hand, strategies may be *randomised* (allowing randomisation over actions) or *pure* (not allowing such randomisation). On the other hand, one can restrict to strategies that use finite memory or even to *stationary* strategies, which only depend on the last state. Indeed, we show that these restrictions give rise to distinct decision problems, which have to be analysed separately.

Our results show that the complexity of NE highly depends on the type of strategies that realise the equilibrium. In particular, we prove the following results, which yield an almost complete picture of the complexity of NE:

1. NE for pure stationary strategies (or pure strategies with bounded memory) is NP-complete.
2. NE for stationary strategies (or randomised strategies with bounded memory) is decidable in PSPACE, but hard for both NP and SqrtSum.
3. NE for arbitrary pure strategies is NP-complete.
4. NE for arbitrary randomised strategies is undecidable.

All of our lower bounds for NE and, in particular, our undecidability result hold already for a subclass of concurrent games where Nash equilibria are guaranteed to exist, namely for *turn-based* games. If this assumption is relaxed and Nash equilibria are not guaranteed to exist, we prove that even the plain existence problem for Nash equilibria is undecidable. Moreover, many of our lower bounds hold already for games where non-zero rewards only occur on terminal states, and thus also for games where each player wants to maximise the *total sum* of the rewards.

As a byproduct of our decidability proof for pure strategies, we give a polynomial-time algorithm for deciding whether in a multi-weighted graph there exists a path whose limit-average weight vector lies between two given thresholds, a result that is of independent interest. For instance, our algorithm can be used for deciding the emptiness of a *multi-threshold mean-payoff language* [2] in polynomial time.

RELATED WORK. Concurrent and, more generally, stochastic games go back to Shapley [24], who proved the existence of the *value* for *discounted two-player zero-sum* games. This result was later generalised by Fink [13] who proved that every discounted game has a Nash equilibrium. Gillette [16] introduced limit-average objectives, and Mertens & Neyman [19] proved the existence of the value for stochastic two-player zero-sum games with limit-average objectives. Unfortunately, as demonstrated by Everett [12], these games do, in general, not admit a Nash equilibrium (see Example 1). However, Vielle [29, 30] proved that, for all $\varepsilon > 0$, every two-player stochastic limit-average game admits an *ε-equilibrium*, i.e. a pair of strategies where each player can gain at most ε from switching her strategy. Whether such equilibria always exist in games with more than two players is an important open question [21].

Determining the complexity of Nash equilibria has attracted much interest in recent years. In particular, a series of papers culminated in the result that computing a Nash equilibrium of a finite two-player game in *strategic form* is complete for the complexity class PPAD [6, 8]. The constrained existence problem, where one looks for a Nash equilibrium with certain properties, has also been investigated for games in strategic form. In particular, Conitzer & Sandholm [7] showed that deciding whether there exists a Nash equilibrium whose payoff exceeds a given threshold and related decision problems are NP-complete for two-player games in strategic form.

For concurrent games with limit-average objectives, most algorithmic results concern two-player zero-sum games. In the turn-based case, these games are commonly known as *mean-payoff games* [10, 32]. While it is known that the value of such a game can be computed in pseudo-polynomial time, it is still open whether there exists a polynomial-time algorithm for solving mean-payoff games. A related model are *multi-dimensional mean-payoff games* where one player tries to maximise several mean-payoff conditions at the same time [5]. In particular, Velner & Rabinovich [28] showed that the value problem for these games is coNP-complete.

One subclass of concurrent games with limit-average objectives that has been studied in the multiplayer setting are concurrent games with reachability objectives. In particular, Bouyer et al. [3] showed that the constrained existence problem for Nash equilibria is NP-complete for these games (see also [26, 14]). We extend their result to limit-average objectives. However, we assume that

strategies can observe actions (a common assumption in game theory), which they do not. Hence, while our result is more general w.r.t. the type of objectives we consider, their result is more general w.r.t. the type of strategies they allow.

In a recent paper [27], we studied the complexity of Nash equilibria in *stochastic* games with reachability objectives. In particular, we proved that NE for pure strategies is undecidable in this setting. Since we prove here that this problem is decidable in the non-stochastic setting, this undecidability result can be explained by the presence of probabilistic transitions in stochastic games. On the other hand, we prove in this paper that randomisation in strategies also leads to undecidability, a question that was left open in [27].

2 Concurrent Games

Concurrent games are played by finitely many players on a finite state space. Formally, a concurrent game is given by

- a finite nonempty set Π of *players*, e.g. $\Pi = \{0, 1, \dots, k-1\}$,
- a finite nonempty set S of *states*,
- for each player i and each state s a nonempty set $\Gamma_i(s)$ of *actions* taken from a finite set Γ,
- a *transition function* $\delta\colon S \times \Gamma^{\Pi} \to S$,
- for each player $i \in \Pi$ a *reward function* $r_i\colon S \to \mathbb{R}$.

For computational purposes, we assume that all rewards are rational numbers with numerator and denominator given in binary. We say that an action profile $\bar{a} = (a_i)_{i\in\Pi}$ is *legal* at state s if $a_i \in \Gamma_i(s)$ for each $i \in \Pi$. Finally, we call a state s *controlled* by player i if $|\Gamma_j(s)| = 1$ for all $j \neq i$, and we say that a game is *turn-based* if each state is controlled by (at least) one player. For turn-based games, an action of the controlling player prescribes to go to a certain state. Hence, we will usually omit actions in turn-based games.

For a tuple $\bar{x} = (x_i)_{i\in\Pi}$, where the elements x_i belong to an arbitrary set X, and an element $x \in X$, we denote by $\bar{x}_{-i}$ the restriction of $\bar{x}$ to $\Pi \setminus \{i\}$ and by $(\bar{x}_{-i}, x)$ the unique tuple $\bar{y} \in X^{\Pi}$ with $y_i = x$ and $\bar{y}_{-i} = \bar{x}_{-i}$.

A play of a game $\mathcal{G}$ is an infinite sequence $s_0\bar{a}_0s_1\bar{a}_1\ldots \in (S \cdot \Gamma^{\Pi})^{\omega}$ such that $\delta(s_j, \bar{a}_j) = s_{j+1}$ for all $j \in \mathbb{N}$. For each player, a play $\pi = s_0\bar{a}_0s_1\bar{a}_1\ldots$ gives rise to an infinite sequence of rewards. There are different criteria to evaluate this sequence and map it to a *payoff*. In this paper, we consider the *limit-average* (or *mean-payoff*) criterion, where the payoff of π for player i is defined by

$$\phi_i(\pi) := \liminf_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} r_i(s_j).$$

Note that this payoff mapping is *prefix-independent*, i.e. $\phi_i(\pi) = \phi_i(\pi')$ if π' is a suffix of π. An important special case are games where non-zero rewards occur only on *terminal* states, i.e. states s with $\delta(s, \bar{a}) = s$ for all (legal) $\bar{a} \in \Gamma^{\Pi}$. These games were introduced by Everett [12] under the name *recursive games*, but we prefer to call them *terminal-reward games*. Hence, in a terminal-reward game, $\phi_i(\pi) = r_i(s)$ if π enters a terminal state s and $\phi_i(\pi) = 0$ otherwise.

Often, it is convenient to designate an *initial* state. An *initialised* game is thus a tuple $(\mathcal{G}, s_0)$ where $\mathcal{G}$ is a concurrent game and s_0 is one of its states.

STRATEGIES AND STRATEGY PROFILES. For a finite set X, we denote by $\mathcal{D}(X)$ the set of probability distributions over X. A *(randomised) strategy* for player i in $\mathcal{G}$ is a mapping $\sigma\colon (S \cdot \Gamma^{\Pi})^* \cdot S \to \mathcal{D}(\Gamma)$ assigning to each possible *history* $xs \in (S \cdot \Gamma^{\Pi})^* \cdot S$ a probability distribution $\sigma(xs)$ over actions such that $\sigma(xs)(a) > 0$ only if $a \in \Gamma_i(s)$. We write $\sigma(a \mid xs)$ for the probability assigned to $a \in \Gamma$ by the distribution $\sigma(xs)$. A *(randomised) strategy profile of* $\mathcal{G}$ is a tuple $\bar{\sigma} = (\sigma_i)_{i \in \Pi}$ of strategies in $\mathcal{G}$, one for each player. Note that a strategy profile can be identified with a function $\bar{\sigma}\colon (S \cdot \Gamma^{\Pi})^* \cdot S \to \mathcal{D}(\Gamma)^{\Pi}$.

A strategy σ for player i is called *pure* if for each $xs \in (S \cdot \Gamma^{\Pi})^* \cdot S$ the distribution $\sigma(xs)$ is *degenerate*, i.e. there exists $a \in \Gamma_i(s)$ with $\sigma(a \mid xs) = 1$. Note that a pure strategy can be identified with a function $\sigma\colon (S \cdot \Gamma^{\Pi})^* \cdot S \to \Gamma$. A strategy profile $\bar{\sigma} = (\sigma_i)_{i \in \Pi}$ is called *pure* if each σ_i is pure, in which case we can identify $\bar{\sigma}$ with a mapping $(S \cdot \Gamma^{\Pi})^* \cdot S \to \Gamma^{\Pi}$. Note that, given an initial state s_0 and a pure strategy profile $\bar{\sigma}$, there exists a unique play $\pi = s_0 \bar{a}_0 s_1 \bar{a}_1 \ldots$ such that $\bar{\sigma}(s_0 \bar{a}_0 \ldots \bar{a}_{j-1} s_j) = \bar{a}_j$ for all $j \in \mathbb{N}$; we call π the play *induced* by $\bar{\sigma}$ from s_0.

A *memory structure* for $\mathcal{G}$ is a triple $\mathfrak{M} = (M, \delta, m_0)$, where M is a set of *memory states*, $\delta\colon M \times S \times \Gamma^{\Pi} \to M$ is the *update function*, and $m_0 \in M$ is the *initial memory*. A (*randomised*) *strategy with memory* $\mathfrak{M}$ for player i is a function $\sigma\colon M \times S \to \mathcal{D}(\Gamma)$ such that $\sigma(m, s)(a) > 0$ only if $a \in \Gamma_i(s)$. The strategy σ is *pure* if the distribution $\sigma(m, s)$ is degenerate for all $m \in M$ and $s \in S$. A (pure) strategy σ with memory $\mathfrak{M}$ can be viewed as a (pure) strategy σ' in the usual sense by setting $\sigma'(xs) = \sigma(\delta^*(x), s)$, where $\delta^*(x)$ is defined inductively by $\delta^*(\varepsilon) = m_0$ and $\delta^*(x \cdot s\bar{a}) = \delta(\delta^*(x), s, \bar{a})$. A *finite-state* strategy is a strategy σ with finite memory $\mathfrak{M}$. If the memory $\mathfrak{M}$ is a singleton, we call σ *stationary*. Moreover, we call a strategy *positional* if it is both pure and stationary. A stationary strategy can thus be represented by a mapping $\sigma\colon S \to \mathcal{D}(\Gamma)$, and a positional strategy by a mapping $\sigma\colon S \to \Gamma$. Finally, we call a strategy profile finite-state, stationary or positional if each strategy in the profile has the respective property.

THE PROBABILITY MEASURE INDUCED BY A STRATEGY PROFILE. Given an initial state $s_0 \in S$ and a strategy profile $\bar{\sigma} = (\sigma_i)_{i \in \Pi}$, the *conditional probability* of

$\bar{a} \in \Gamma^{\Pi}$ given the history $xs \in (S \cdot \Gamma^{\Pi})^* \cdot S$ equals

$$\bar{\sigma}(\bar{a} \mid xs) := \prod_{i \in \Pi} \sigma_i(a_i \mid xs).$$

The probabilities $\bar{\sigma}(\bar{a} \mid xs)$ induce a probability measure on the Borel σ-algebra over $(S \cdot \Gamma^{\Pi})^{\omega}$ as follows: The probability of a basic open set $s_1\bar{a}_1 \dots s_n\bar{a}_n \cdot (S \cdot \Gamma^{\Pi})^{\omega}$ equals the product $\prod_{j=1}^{n} \bar{\sigma}(\bar{a}_j \mid s_1\bar{a}_1 \dots \bar{a}_{j-1}s_j)$ if $s_1 = s_0$ and $\delta(s_j, \bar{a}_j) = s_{j+1}$ for all $1 \leq j < n$; in all other cases, this probability is 0. By *Carathéodory's extension theorem*, this extends to a unique probability measure assigning a probability to every Borel subset of $(S \cdot \Gamma^{\Pi})^{\omega}$, which we denote by $\mathrm{Pr}^{\bar{\sigma}}_{s_0}$. Via the natural projection $(S \cdot \Gamma^{\Pi})^{\omega} \to S^{\omega}$, we obtain a probability measure on the Borel σ-algebra over S^{ω}. We abuse notation and denote this measure also by $\mathrm{Pr}^{\bar{\sigma}}_{s_0}$; it should always be clear from the context to which measure we are referring to. Finally, we denote by $\mathrm{E}^{\bar{\sigma}}_{s_0}$ the expectation operator that corresponds to $\mathrm{Pr}^{\bar{\sigma}}_{s_0}$, i.e. $\mathrm{E}^{\bar{\sigma}}_{s_0}(f) = \int f \, \mathrm{d}\mathrm{Pr}^{\bar{\sigma}}_{s_0}$ for all Borel measurable functions $f \colon (S \cdot \Gamma^{\Pi})^{\omega} \to \mathbb{R} \cup \{\pm\infty\}$ or $f \colon S^{\omega} \to \mathbb{R} \cup \{\pm\infty\}$. In particular, we are interested in the quantities $p_i := \mathrm{E}^{\bar{\sigma}}_{s_0}(\phi_i)$. We call p_i the *(expected) payoff* of $\bar{\sigma}$ for player i and the vector $(p_i)_{i \in \Pi}$ the (expected) payoff of $\bar{\sigma}$. Finally, we call a history $x \in (S \cdot \Gamma^{\Pi})^* \cdot S$ *consistent* with $\bar{\sigma}$ if $\mathrm{Pr}^{\bar{\sigma}}_{s_0}(x \cdot (S \cdot \Gamma^{\Pi})^{\omega}) > 0$.

In order to apply known results about Markov chains, we can also view the stochastic process induced by a strategy profile $\bar{\sigma}$ as a countable Markov chain $\mathcal{G}^{\bar{\sigma}}$, defined as follows: The set of states of $\mathcal{G}^{\bar{\sigma}}$ equals the set $(S \cdot \Gamma^{\Pi})^* \cdot S$ of histories of $\mathcal{G}$. The only transitions from a state xs lead to states of the form $xs\bar{a}t$ where $t = \delta(s, \bar{a})$, and such a transition occurs with probability $\bar{\sigma}(\bar{a} \mid xs)$.

For each player i, the Markov decision process $\mathcal{G}^{\bar{\sigma}_{-i}}$ has the same states as $\mathcal{G}^{\bar{\sigma}}$, and there is a transition from a state xs to a state $xs\bar{a}t$ with action $a \in \Gamma_i(s)$ and probability p if $a_i = a$, $\delta(s, \bar{a}) = t$ and $p = \prod_{j \neq i} \sigma_j(a_j)$. Finally, the reward of a state xs in $\mathcal{G}^{\bar{\sigma}_{-i}}$ equals the reward $r_i(s)$ of the state s for player i in $\mathcal{G}$.

If $\bar{\sigma}$ is a strategy profile with finite memory $\mathfrak{M}$, we make $\mathcal{G}^{\bar{\sigma}}$ and $\mathcal{G}^{\bar{\sigma}_{-i}}$ finite by quotienting the state space w.r.t. the equivalence relation $\sim$, defined by $xs \sim yt$ if $s = t$ and $\delta^*(x) = \delta^*(y)$. In particular, if $\bar{\sigma}$ is stationary, then the state spaces of $\mathcal{G}^{\bar{\sigma}}$ and $\mathcal{G}^{\bar{\sigma}_{-i}}$ coincide with the state space of $\mathcal{G}$.

DRAWING CONCURRENT GAMES. When drawing a concurrent game as a graph, we will adhere to the following conventions: States are usually depicted as circles, but terminal states are depicted as squares. The initial state is marked by a dangling incoming edge. An edge from s to t with label $\bar{a}$ means that $\delta(s, \bar{a}) = t$ and that $\bar{a}$ is legal at s. However, the label $\bar{a}$ might be omitted if it is not essential. In turn-based games, the player who controls a state is indicated by the label next to it. Finally, a label of the form $i \colon x$ next to state s indicates that $r_i(s) = x$; if this reward is 0, the label will usually be omitted.

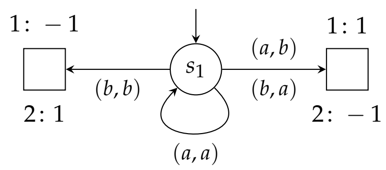

Figure 1. A terminal-reward game that has no Nash equilibrium.

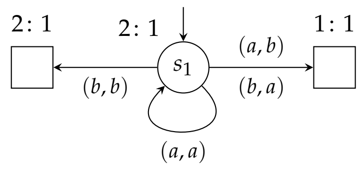

Figure 2. A limit-average game that has no Nash equilibrium.

3 Nash Equilibria

To capture rational behaviour of selfish players, Nash [20] introduced the notion of — what is now called — a *Nash equilibrium*. Formally, given a game $\mathcal{G}$ and an initial state s_0, a strategy τ for player i is a *best response* to a strategy profile $\bar{\sigma}$ if τ maximises the expected payoff for player i, i.e.

$$\mathrm{E}_{s_0}^{\bar{\sigma}_{-i},\tau'}(\phi_i) \leq \mathrm{E}_{s_0}^{\bar{\sigma}_{-i},\tau}(\phi_i)$$

for all strategies τ' for player i. A strategy profile $\bar{\sigma} = (\sigma_i)_{i \in \Pi}$ is a *Nash equilibrium* of $(\mathcal{G}, s_0)$ if for each player i the strategy σ_i is a best response to $\bar{\sigma}$. Hence, in a Nash equilibrium no player can improve her payoff by (unilaterally) switching to a different strategy. As the following examples demonstrate, Nash equilibria are not guaranteed to exist in concurrent games.

Example 1. Consider the terminal-reward game $\mathcal{G}_1$ depicted in Figure 1 and played by players 1 and 2, which was originally presented in [9]. We claim that $(\mathcal{G}_1, s_1)$ does not have a Nash equilibrium. First note that, for each $\varepsilon > 0$, player 1 can ensure a payoff of $1 - 2\varepsilon$ by the stationary strategy that selects action b with probability ε. Hence, every Nash equilibrium (σ, τ) of $(\mathcal{G}_1, s_1)$ must have payoff $(1, -1)$. Now we distinguish whether $\sigma(b \mid (s_1(a,a))^k s_1) = 0$ for all $k \in \mathbb{N}$ or not. In the first case, there must exist $k \in \mathbb{N}$ such that $\tau(b \mid (s_1(a,a))^k s_1) > 0$ (otherwise (σ, τ) would not have payoff $(1, -1)$). But then Player 2 can improve her payoff by always playing action a with probability 1, a contradiction to (σ, τ) being a Nash equilibrium. In the second case, consider the least k such that $p := \sigma(b \mid (s_1(a,a))^k s_1) > 0$. By choosing action b with probability 1 for the history $(s_1(a,a))^k s_1$ and choosing action a with probability 1 for all other histories, player 2 can ensure payoff p, again a contradiction to (σ, τ) being a Nash equilibrium.

Example 2. A variation of the previous game is the game $\mathcal{G}_2$, which is depicted in Figure 2 and also played by players 1 and 2. It is not a terminal-reward game, but the only rewards that occur in the game are 0 and 1. Using almost the same argumentation as in Example 1, we can show that $(\mathcal{G}_2, s_1)$ has no Nash equilibrium either.

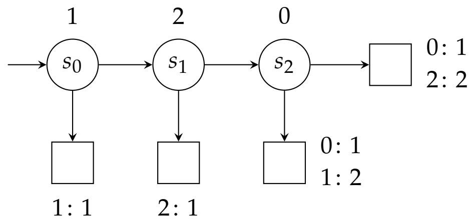

Figure 3. A game with no pure Nash equilibrium where player 0 wins with positive probability.

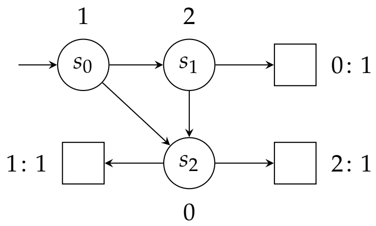

Figure 4. A game with no stationary Nash equilibrium where player 0 wins with positive probability.

It follows from Nash's theorem [20] that every game whose arena is a tree (or a DAG) has a Nash equilibrium. Another important special case of concurrent limit-average games where Nash equilibria always exist are turn-based games. For these games, Thuijsman & Raghavan [25] proved not only the existence of arbitrary Nash equilibria but of pure finite-state ones.

To measure the complexity of Nash equilibria in concurrent games, we introduce the following decision problem, which we call NE:

> Given a game $\mathcal{G}$, a state s_0 and thresholds $\bar{x}, \bar{y} \in (\mathbb{Q} \cup \{\pm\infty\})^{\Pi}$, decide whether $(\mathcal{G}, s_0)$ has a Nash equilibrium with payoff $\geq \bar{x}$ and $\leq \bar{y}$.

Note that we have not put any restriction on the type of strategies that realise the equilibrium. It is natural to restrict the search space to profiles of pure, stationary or positional strategies. These restrictions give rise to different decision problems, which we call PureNE, StatNE and PosNE, respectively.

Before we analyse the complexity of these problems, let us convince ourselves that these problems are not just different faces of the same coin. We first show that the decision problems where we look for equilibria in randomised strategies are distinct from the ones where we look for equilibria in pure strategies.

Proposition 3. There exists a turn-based terminal-reward game that has a stationary Nash equilibrium where player 0 receives payoff 1 but that has no pure Nash equilibrium where player 0 receives payoff > 0.

Proof. Consider the game depicted in Figure 3 and played by three players 0, 1 and 2. Clearly, the stationary strategy profile where at state s_2 player 0 selects both outgoing transitions with probability $\frac{1}{2}$ each, player 1 plays from s_0 to s_1 and player 2 plays from s_1 to s_2 is a Nash equilibrium where player 0 receives payoff 1. However, in any pure strategy profile where player 0 receives payoff > 0, either player 1 or player 2 receives payoff 0 and could improve her payoff by switching her strategy at s_0 or s_1, respectively. $\square$

Now we show that it makes a difference whether we look for an equilibrium in stationary strategies or not.

Proposition 4. There exists a turn-based terminal-reward game that has a pure Nash equilibrium where player 0 receives payoff 1 but that has no stationary Nash equilibrium where player 0 receives payoff > 0.

Proof. Consider the game $\mathcal{G}$ depicted in Figure 4 and played by three players 0, 1 and 2. Clearly, the pure strategy profile that leads to the terminal state with payoff 1 for player 0 and where player 0 plays "right" if player 1 has deviated and "left" if player 2 has deviated is a Nash equilibrium of $(\mathcal{G}, s_0)$ with payoff 1 for player 0. Now consider any stationary equilibrium of $(\mathcal{G}, s_0)$ where player 0 receives payoff > 0. If the stationary strategy of player 0 prescribes to play "right" with positive probability, then player 2 can improve her payoff by playing to s_2 with probability 1, and otherwise player 1 can improve her payoff by playing to s_2 with probability 1, a contradiction. □

It follows from Section 3 that NE and StatNE are different from PureNE and PosNE, and it follows from Proposition 4 that NE and PureNE are different from StatNE and PosNE. Hence, all of these decision problems are pairwise distinct, and their decidability and complexity has to be studied separately.

4 Positional Strategies

In this section, we show that the problem PosNE is NP-complete; we start by proving the upper bound.

Theorem 5. PosNE is in NP.

Proof. To decide PosNE on input $\mathcal{G}, s_0, \bar{x}, \bar{y}$, we start by guessing a positional strategy profile $\bar{\sigma}$ of $\mathcal{G}$, i.e. mappings $\sigma_i \colon S \to \Gamma$ such that $\sigma_i(s) \in \Gamma_i(s)$ for all $i \in \Pi$ and $s \in S$. Then, we verify whether $\bar{\sigma}$ is a Nash equilibrium with the desired payoff. To do this, we first compute the payoff z_i of $\bar{\sigma}$ for each player i by computing the number $\mathrm{E}^{\bar{\sigma}}_{s_0}(\phi_i)$ in the finite Markov chain $\mathcal{G}^{\bar{\sigma}}$. Since $\mathcal{G}^{\bar{\sigma}}$ is deterministic, this number equals the average weight (for player i) on the unique simple cycle reachable from s_0 and can thus be computed in polynomial time. Once each z_i is computed, we can easily check whether $x_i \leq z_i \leq y_i$. To verify that $\bar{\sigma}$ is a Nash equilibrium, we additionally compute, for each player i, the value v_i of the finite MDP $\mathcal{G}^{\bar{\sigma}_{-i}}$ from s_0. This number can be computed by identifying the highest average weight (for player i) on a simple cycle reachable in $\mathcal{G}^{\bar{\sigma}_{-i}}$ from s_0, which can also be done in polynomial time [18]. Clearly, $\bar{\sigma}$ is a Nash equilibrium if and only if $v_i \leq z_i$ for each player i. □

A result by Chatterjee et al. [5, Lemma 15] implies that PosNE is NP-hard, even for turn-based games with rewards taken from $\{-1, 0, 1\}$ (but with an unbounded number of players). We strengthen their result by showing that the problem remains NP-hard if there are only three players and rewards are taken from $\{0, 1\}$.

Theorem 6. PosNE is NP-hard, even for turn-based three-player games with rewards 0 and 1.

Proof. We reduce from the Hamiltonian cycle problem. Given a graph $G = (V, E)$, we define a turn-based three-player game $\mathcal{G}$ as follows: the set of states is V, all states are controlled by player 0, and the transition function corresponds to E (i.e. $\Gamma_0(v) = vE$ and $\delta(v, \bar{a}) = w$ if and only if $a_0 = w$). Let $n = |V|$ and $v_0 \in V$. Player 0 receives reward 1 in each state. The reward of state v_0 to player 1 equals 1; all other states have reward 0 for player 1. Finally, player 2 receives reward 0 at v_0 and reward 1 at all other states. We show that there is a Hamiltonian cycle in G if and only if $(\mathcal{G}, v_0)$ has a positional Nash equilibrium with payoff $\geq (1, 1/n, (n-1)/n)$.

($\Rightarrow$) Let $\pi = \pi(0)\pi(1)\ldots\pi(n)$ be a Hamiltonian cycle that starts (and ends) in $\pi(0) = v_0 = \pi(n)$. Consider the positional strategy σ of player 0 that plays from $\pi(i)$ to $\pi(i+1)$ for all $i < n$. The induced play from v_0 is the play $(\pi(0)\pi(1)\ldots\pi(n-1))^\omega$, which gives payoff 1 to player 0, payoff $1/n$ to player 1 and payoff $(n-1)/n$ to player 2. Moreover, it is obvious that we have a Nash equilibrium.

($\Leftarrow$) Let π be the play induced by a positional Nash equilibrium of $(\mathcal{G}, v_0)$ with payoff $\geq (1, 1/n, (n-1)/n)$. Since π corresponds to a positional strategy profile and gives player 1 a positive payoff, π has the form $\pi = (v_0 v_1 \ldots v_{i-1})^\omega$, where $1 \leq i \leq n$ and $v_0 \ldots v_{i-1} v_0$ is a simple cycle of G. Hence, the payoff of π for player 2 equals $(i-1)/i$. This number is greater than $(n-1)/n$ only if $i \geq n$. Hence, $i = n$ and $v_0 \ldots v_{i-1} v_0$ is a Hamiltonian cycle. $\square$

By combining our reduction with a game that has no positional Nash equilibrium, we can prove the following stronger result for non-turn-based games.

Corollary 7. Deciding the existence of a positional Nash equilibrium in a concurrent limit-average game is NP-complete, even for three-player games with rewards 0 and 1.

Proof. Membership in NP follows from Theorem 5. To prove hardness, we reduce from the following problem, whose NP-hardness follows from the proof of Theorem 6: Given a three-player game $(\mathcal{G}, s_0)$ with rewards 0 and 1 and $n \in \mathbb{N}$ (given in unary), decide whether $(\mathcal{G}, s_0)$ has a positional Nash equilibrium with payoff $\geq (1, 1/n, (n-1)/n)$. From $\mathcal{G}$, we construct a new game $\mathcal{G}'$, which employs the game $\mathcal{G}_2$ from Example 2 and is depicted in Figure 5; we set the reward for player 0 in all states of $\mathcal{G}_2$ to 1. Note that we can simulate the fractional rewards in the terminal state by a cycle of n states with rewards 0 and 1. We claim that $(\mathcal{G}', s_0')$ has a positional Nash equilibrium if and only if $(\mathcal{G}, s_0)$ has a positional Nash equilibrium with payoff $\geq (1, 1/n, (n-1)/n)$.

($\Rightarrow$) Let $\bar{\sigma}$ be a positional Nash equilibrium of $(\mathcal{G}', s_0')$. Since $(\mathcal{G}_2, s_1)$ does not have a Nash equilibrium, the induced play must either enter the game $\mathcal{G}$

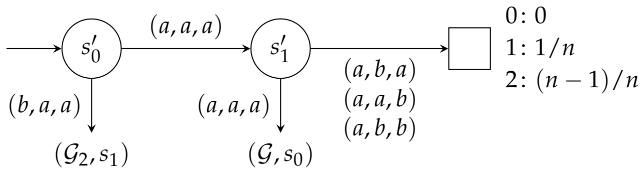

Figure 5. The game $\mathcal{G}'$.

or end at the terminal state with payoff 0 for player 0. But the latter case is impossible since then player 0 could improve her payoff by playing action b at s'_0. Hence, the induced play enters $\mathcal{G}$, and $\bar{\sigma}$ is also a Nash equilibrium of $(\mathcal{G}, s_0)$. Moreover, $\bar{\sigma}$ must have payoff at least $(1, 1/n, (n-1)/n)$ since otherwise player 1 or player 2 could improve her payoff by playing action b at s'_1.

($\Leftarrow$) Let $\bar{\sigma}$ be a positional Nash equilibrium of $(\mathcal{G}, s_0)$ with payoff at least $(1, 1/n, (n-1)/n)$. We can extend $\bar{\sigma}$ to a positional Nash equilibrium of $(\mathcal{G}', s'_0)$ by setting $\bar{\sigma}(s'_0) = \bar{\sigma}(s'_0(a,a,a)s'_1) = (a,a,a)$. □

5 Stationary Strategies

To prove the decidability of StatNE, we appeal to results established for the *existential theory of the reals*, the set of all existential first-order sentences (over the appropriate signature) that hold in the ordered field $\mathfrak{R} := (\mathbb{R}, +, \cdot, 0, 1, \leq)$. The best known upper bound for the complexity of the associated decision problem is PSPACE [4], which leads to the following theorem.

Theorem 8. StatNE is in PSPACE.

Proof. To prove membership in PSPACE, we show that there is a polynomial-time procedure that on input $\mathcal{G}, s_0, \bar{x}, \bar{y}$ returns an existential first-order sentence ψ such that $(\mathcal{G}, s_0)$ has a stationary Nash equilibrium with payoff $\geq \bar{x}$ and $\leq \bar{y}$ if and only if ψ holds in $\mathfrak{R}$. How does ψ look like? Let $\bar{\alpha} = (\alpha^i_{s,a})_{i\in\Pi, s\in S, a\in\Gamma}$, $\bar{v} = (v^i_s)_{i\in\Pi, s\in S}$, $\bar{b} = (b_s)_{s\in S}$ and $\bar{z} = (z^i_s)_{i\in\Pi, s\in S}$ be four sets of variables. The formula

$$\varphi_i(\bar{\alpha}) := \bigwedge_{s\in S} \Big(\sum_{a\in\Gamma} \alpha^i_{s,a} = 1 \wedge \bigwedge_{a\in\Gamma_i(s)} \alpha^i_{s,a} \geq 0 \wedge \bigwedge_{a\in\Gamma\setminus\Gamma_i(s)} \alpha^i_{s,a} = 0 \Big)$$

states that the mapping $\sigma_i \colon S \to \mathbb{R}^\Gamma$, defined by $\sigma_i(s) \colon a \mapsto \alpha^i_{s,a}$ is indeed a stationary strategy for player i. Provided that each $\varphi_i(\bar{\alpha})$ holds in $\mathfrak{R}$, the formula

$$\eta_i(\bar{\alpha}, \bar{z}) := \exists \bar{b} \Big(\bigwedge_{s\in S} b_s + z^i_s = r_i(s) + \sum_{\bar{a}\in\Gamma^\Pi} b_{\delta(s,\bar{a})} \cdot \prod_{j\in\Pi} \alpha^j_{s,a_j} \Big) \wedge$$
$$\bigwedge_{s\in S} z^i_s = \sum_{\bar{a}\in\Gamma^\Pi} z^i_{\delta(s,\bar{a})} \cdot \prod_{j\in\Pi} \alpha^j_{s,a_j}$$

states that $z_s^i = \mathrm{E}_s^{\bar{\sigma}}(\phi_i)$ for all $s \in S$, where $\bar{\sigma} = (\sigma_i)_{i \in \Pi}$ (see [22, Theorem 8.2.6]). Finally, the formula

$$\vartheta_i(\bar{\alpha}, \bar{v}) := \exists \bar{b} \Big(\bigwedge_{s \in S} \bigwedge_{a \in \Gamma} b_s + v_s^i \geq r_i(s) + \sum_{\substack{\bar{a} \in \Gamma^\Pi \\ a_i = a}} b_{\delta(s,\bar{a})} \cdot \prod_{j \neq i} \alpha_{s,a_j}^j \Big) \wedge \\ \bigwedge_{s \in S} \bigwedge_{a \in \Gamma} v_s^i \geq \sum_{\substack{\bar{a} \in \Gamma^\Pi \\ a_i = a}} v_{\delta(s,\bar{a})}^i \cdot \prod_{j \neq i} \alpha_{s,a_j}^j$$

states that $\bar{v}$ is a solution of the linear programme for computing the values of the MDP $\mathcal{G}^{\bar{\sigma}_{-i}}$ (see [22, Section 9.3]), i.e. the formula is fulfilled if and only if $v_s^i \geq \sup_\tau \mathrm{E}_s^{\bar{\sigma}_{-i},\tau}(\phi_i)$ for all $i \in \Pi$ and $s \in S$.

The desired sentence ψ is the existential closure of the conjunction of the formulae φ_i, η_i and ϑ_i combined with formulae stating that player i cannot improve her payoff and that the expected payoff for player i lies in-between the given thresholds:

$$\psi := \exists \bar{\alpha}\, \exists \bar{v}\, \exists \bar{z} \bigwedge_{i \in \Pi} (\varphi_i(\bar{\alpha}) \wedge \eta_i(\bar{\alpha}, \bar{z}) \wedge \vartheta_i(\bar{\alpha}, \bar{v}) \wedge v_{s_0}^i \leq z_{s_0}^i \wedge x_i \leq z_{s_0}^i \leq y_i)\,.$$

Clearly, ψ can be constructed in polynomial time from $\mathcal{G}$, s_0, $\bar{x}$ and $\bar{y}$. Moreover, ψ holds in $\mathfrak{R}$ if and only if $(\mathcal{G}, s_0)$ has a stationary Nash equilibrium with payoff at least $\bar{x}$ and at most $\bar{y}$. $\square$

The next theorem shows that StatNE is NP-hard, even for turn-based games with rewards 0 and 1. Note that this does not follow from the NP-hardness of PosNE, but requires a different proof.

Theorem 9. StatNE is NP-hard, even for turn-based games with rewards 0 and 1.

Proof. We employ a reduction from SAT, which resembles a reduction in [26]. Given a Boolean formula $\varphi = C_1 \wedge \cdots \wedge C_m$ in conjunctive normal form over propositional variables $X_1, \ldots, X_n$, where w.l.o.g. $m \geq 1$ and each clause is nonempty, we build a turn-based game $\mathcal{G}$ played by players $0, 1, \ldots, n$ as follows: The game $\mathcal{G}$ has states $C_1, \ldots, C_m$ controlled by player 0 and for each clause C and each literal L that occurs in C a state (C, L), controlled by player i if $L = X_i$ or $L = \neg X_i$; additionally, the game contains a terminal state $\bot$. There are transitions from a clause C_j to each state (C_j, L) such that L occurs in C_j and from there to $C_{(j \bmod m)+1}$, and there is a transition from each state of the form $(C, \neg X)$ to $\bot$. Each state except $\bot$ has reward 1 for player 0, whereas $\bot$ has reward 0 for player 0. For player i, all states except states of the form (C, X_i) have reward 1; states of the form (C, X_i) have reward 0. The structure of $\mathcal{G}$ is depicted in Figure 6.

Clearly, $\mathcal{G}$ can be constructed from φ in polynomial time. In order to establish our reduction, we prove that the following statements are equivalent:

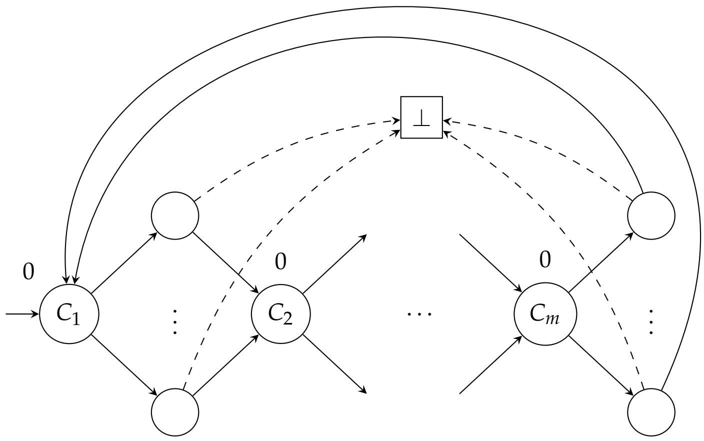

Figure 6. Reducing SAT to StatNE.

1. φ is satisfiable.
2. $(\mathcal{G}, C_1)$ has a positional Nash equilibrium with payoff ≥ 1 for player 0.
3. $(\mathcal{G}, C_1)$ has a stationary Nash equilibrium with payoff ≥ 1 for player 0.

($1. \Rightarrow 2.$) Assume that $\alpha\colon \{X_1, \ldots, X_n\} \to \{\text{true}, \text{false}\}$ is a satisfying assignment for φ. We show that the positional strategy profile $\overline{\sigma}$ where at any time player 0 plays from a clause C to a fixed state (C, L) such that L is mapped to true by α and each player $i \neq 0$ never plays to $\perp$ is a Nash equilibrium of $(\mathcal{G}, C_1)$ with payoff 1 for player 0. First note that the induced play never reaches $\perp$. Hence, player 0 receives payoff 1, which is the best payoff player 0 can get.

To show that $\overline{\sigma}$ is a Nash equilibrium, consider any player $i \neq 0$ who receives payoff < 1. Hence, a state of the form (C, X_i) is visited in the induced play. However, as player 0 plays according to the satisfying assignment, no state of the form $(C', \neg X_i)$ is ever visited. Hence, player i cannot improve her payoff by playing to $\perp$.

($2. \Rightarrow 3.$) Trivial.

($3. \Rightarrow 1.$) Assume that $(\mathcal{G}, C_1)$ has a stationary Nash equilibrium $\overline{\sigma}$ with payoff ≥ 1 for player 0. Hence, the terminal state $\perp$ is reached with probability 0 in $\overline{\sigma}$. Consider the variable assignment α that maps X_i to true if and only if player i receives payoff < 1 from $\overline{\sigma}$; we claim that α satisfies the formula. Consider any clause C. By the construction of $\mathcal{G}$, there exists a literal $L \in C$ such that $\sigma_0((C, L) \mid C) > 0$. If $L = X_i$, then $\mathrm{E}^{\overline{\sigma}}_{C_1}(\phi_i) < 1$ and α maps X_i to true, thus satisfying C. If $L = \neg X_i$, then player i must receive payoff 1 since otherwise she could switch to the positional strategy τ that plays from (C, L) to $\perp$; in the strategy profile $(\overline{\sigma}_{-i}, \tau)$ the state $\perp$ is visited with probability 1, which gives payoff 1 to player i. Hence, α maps X_i to false and satisfies C. $\square$

By combining our reduction with the game from Example 1, we can prove the following stronger result for concurrent games.

Corollary 10. Deciding the existence of a stationary Nash equilibrium in a concurrent limit-average game with rewards 0 and 1 is NP-hard.

Proof. The proof is similar to the proof of Corollary 7. From a given concurrent limit-average game $(\mathcal{G}, s_0)$ with rewards 0 and 1, we construct a new game $(\mathcal{G}', s'_0)$ such that $(\mathcal{G}', s'_0)$ has a stationary Nash equilibrium if and only if $(\mathcal{G}, s_0)$ has a stationary Nash equilibrium with payoff at least 1 for player 0. The game $\mathcal{G}'$ is the disjoint union of $\mathcal{G}$, the game $\mathcal{G}_2$ from Example 2, and the state s'_0, which is controlled by player 0. At s'_0 player 0 can either play to the initial state s_0 of $\mathcal{G}$ or to the initial state s_1 of $\mathcal{G}_2$. Finally, we set the reward for player 0 in every state of $\mathcal{G}_2$ to 1. $\square$

So far we have shown that StatNE is contained in PSPACE and hard for NP, leaving a considerable gap between the two bounds. In order to gain a better understanding of StatNE, we relate this problem to the *square root sum problem* (SqrtSum), an important problem about numerical computations. Formally, SqrtSum is the following decision problem: Given numbers $d_1, \ldots, d_n, k \in \mathbb{N}$, decide whether $\sum_{i=1}^{n} \sqrt{d_i} \geq k$. Recently, Allender et al. [1] showed that SqrtSum belongs to the fourth level of the *counting hierarchy*, a slight improvement over the previously known PSPACE upper bound. However, it has been an open question since the 1970s as to whether SqrtSum falls into the polynomial hierarchy [15, 11]. We give a polynomial-time reduction from SqrtSum to StatNE for turn-based terminal-reward games. Hence, StatNE is at least as hard as SqrtSum, and showing that StatNE resides inside the polynomial hierarchy would imply a major breakthrough in understanding the complexity of numerical computations. While our reduction is similar to the one in [27], it requires new techniques to simulate stochastic states.

Theorem 11. SqrtSum is polynomial-time reducible to StatNE for turn-based 8-player terminal-reward games.

Before we state the reduction, let us first examine the game $\mathcal{G}(p)$, where $0 \leq p \leq 1$, which is played by players $0, 1, \ldots, 5$ and depicted in Figure 7.

Lemma 12. The maximal payoff player 1 receives in a stationary Nash equilibrium of $(\mathcal{G}(p), s_1)$ where player 0 receives payoff ≥ 0 equals $\sqrt{p}$.

Proof. Let $\bar{\sigma}$ be a stationary strategy profile of $(\mathcal{G}(p), s_1)$ where player 0 receives payoff ≥ 0, and let $q_i = \sigma_0(v_i \mid u_i)$ be the probability that player 0 moves from u_i to v_i. We claim that $q := q_1 = q_2 = 1 - p$ if $\bar{\sigma}$ is a Nash equilibrium. Let $z = \mathrm{E}_{v_1}^{\bar{\sigma}}(\phi_4)$ and $z' = \mathrm{E}_{v_1}^{\bar{\sigma}}(\phi_5)$. Since $\bar{\sigma}$ is a Nash equilibrium, we have $z \geq 1 - p$ and $z' \geq 1$ (otherwise player 4 or player 5 would prefer to leave the game at r_2

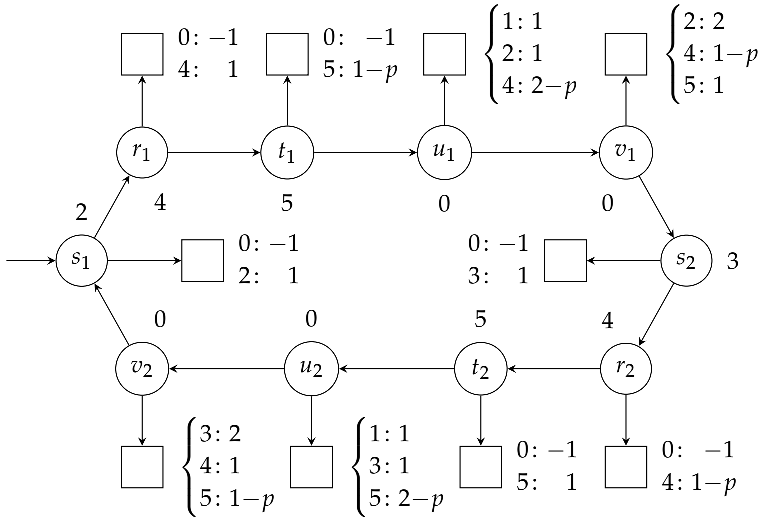

Figure 7. The game $\mathcal{G}(p)$.

or t_2). On the other hand, since at every terminal state the sum of the rewards for players 4 and 5 is at most $2-p$, we have $z+z' \leq 2-p$. Hence, $z = 1-p$ and $z' = 1$. Now consider the expected payoffs for players 4 and 5 from r_1:

$$\mathrm{E}^{\bar{\sigma}}_{r_1}(\phi_4) = (1-q_1)(2-p) + q_1 \cdot z = 2 - q_1 - p;$$
$$\mathrm{E}^{\bar{\sigma}}_{r_1}(\phi_5) = q_1 \cdot z' = q_1\,.$$

Since $\bar{\sigma}$ is a Nash equilibrium, these numbers are bounded from below by 1 and $1-p$, respectively (otherwise, player 4 or player 5 would leave the game at r_1 or t_1). Hence, $q_1 = 1-p$. The reasoning that $q_2 = 1-p$ is analogous.

In the following, assume without loss of generality that $0 < p < 1$ (otherwise the statement of the lemma is trivial). For any stationary strategy profile $\bar{\sigma}$ of $\mathcal{G}(p)$ where player 0 receives payoff ≥ 0, let $x_1 = \sigma_0(s_2 \mid v_1)$ and $x_2 = \sigma_0(s_1 \mid v_2)$ be the probabilities that player 0 does not leave the game at v_1, respectively v_2. Given x_1 and x_2, for $i = 1,2$ we can compute the payoff $f_i(x_1, x_2) := \mathrm{E}^{\bar{\sigma}}_{s_i}(\phi_{i+1})$ for player $i+1$ from s_i by

$$f_i(x_1, x_2) = \frac{p + 2q(1-x_i)}{1 - q^2 x_1 x_2}\,.$$

To have a Nash equilibrium, it must be the case that $f_1(x_1, x_2), f_2(x_1, x_2) \geq 1$ since otherwise player 2 or player 3 would prefer to leave the game at s_1 or s_2, respectively, which would give the respective player payoff 1 immediately. Vice versa, if $f_1(x_1, x_2), f_2(x_1, x_2) \geq 1$ then $\bar{\sigma}$ is a Nash equilibrium with expected payoff

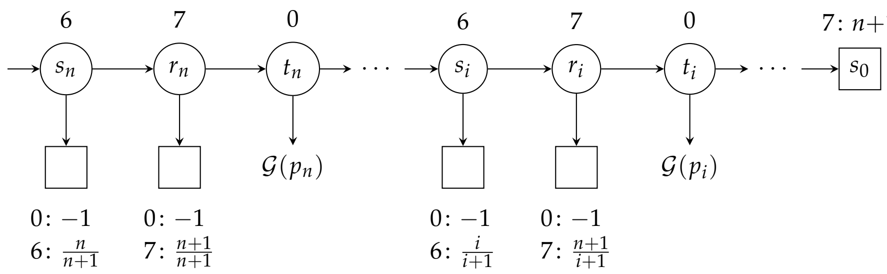

Figure 8. Reducing SqrtSum to StatNE.

$$f(x_1, x_2) := \frac{p + qx_1 p}{1 - q^2 x_1 x_2}$$

for player 1. Hence, to determine the maximum payoff for player 1 in a stationary Nash equilibrium where player 0 receives payoff ≥ 0, we have to maximise $f(x_1, x_2)$ under the constraints $f_1(x_1, x_2), f_2(x_1, x_2) \geq 1$ and $0 \leq x_1, x_2 \leq 1$. We claim that the maximum is reached only if $x_1 = x_2$. If e.g. $x_1 > x_2$, then we can achieve a higher payoff for player 1 by setting $x_2' := x_1$, and the constraints are still satisfied:

$$\frac{p + 2q(1 - x_2')}{1 - q^2 x_1 x_2'} = \frac{p + 2q(1 - x_1)}{1 - q^2 x_1^2} \geq \frac{p + 2q(1 - x_1)}{1 - q^2 x_1 x_2} \geq 1\,.$$

Hence, it suffices to maximise $f(x, x)$ subject to $f_1(x, x) \geq 1$ and $0 \leq x \leq 1$, which is equivalent to maximising $f(x, x)$ subject to $(1 - p)x^2 - 2x + 1 \geq 0$ and $0 \leq x \leq 1$. and The roots of the quadratic function are $(1 \pm \sqrt{p})/(1 - p)$, but $(1 + \sqrt{p})/(1 - p) > 1$ for $p > 0$. Therefore, any solution x must satisfy $x \leq x_0 := (1 - \sqrt{p})/(1 - p)$. Since $0 \leq x_0 \leq 1$ for $0 < p < 1$ and $f(x, x)$ is strictly increasing on $[0, 1]$, the optimal solution is x_0, and the maximal payoff for player 1 in a stationary Nash equilibrium of $(\mathcal{G}(p), s_1)$ where player 0 receives payoff ≥ 0 equals indeed

$$f(x_0, x_0) = \frac{p + qx_0 p}{1 - q^2 x_0^2} = \frac{p}{1 - qx_0} = \frac{p}{1 - (1 - p)x_0} = \frac{p}{1 - (1 - \sqrt{p})} = \sqrt{p}\,. \quad \square$$

Proof (of Theorem 11). Given an instance $(d_1, \ldots, d_n, k)$ of SqrtSum, where w.l.o.g. $n > 0$, $d_i > 0$ for each $i = 1, \ldots, n$, and $d := \sum_{i=1}^{n} d_i$, we construct a turn-based 8-player terminal-reward game $(\mathcal{G}, s)$ such that $(\mathcal{G}, s)$ has a stationary Nash equilibrium with payoff $\geq (0, \frac{k}{d(n+1)}, 0, \ldots, 0)$ if and only if $\sum_{i=1}^{n} \sqrt{d_i} \geq k$. Define $p_i := d_i/d^2$ for $i = 1, \ldots, n$. For the reduction, we use n copies of the game $\mathcal{G}(p)$, where in the ith copy we set p to p_i; in each copy, we set the rewards to player 6 and player 7 at all terminal states to 1 and 0, respectively. The complete game $\mathcal{G}$ is depicted in Figure 8; it can obviously be constructed in polynomial time. We claim that in any (stationary) Nash equilibrium of $(\mathcal{G}, s_n)$ where player 0

receives payoff ≥ 0 the probability of reaching the game $\mathcal{G}(p_i)$ equals $1/(n+1)$ for all $i = 1, \ldots, n$. First note that in any such equilibrium the state s_0 must be reached with positive probability since otherwise player 7 would prefer to leave the game at one of the states r_i, giving player 0 payoff < 0. Now let $\overline{\sigma}$ be a stationary Nash equilibrium of $(\mathcal{G}, s_n)$ where player 0 receives payoff ≥ 0, and set $q_i := \sigma_0(s_{i-1} \mid t_i)$. By induction on i, we prove that $q_i = i/(i+1)$. For $i = 1$, this is true because if $q_i > \frac{1}{2}$ then player 6 would prefer to leave the game at s_1, and if $q_i < \frac{1}{2}$ then player 7 would prefer to leave the game at r_1. Now let $i > 1$ and assume that $q_j = j/(j+1)$ for all $j < i$. A simple calculation reveals that the expected payoffs for player 6 and player 7 from s_{i-1} equal $(i-1)/i$ and $(n+1)/i$, respectively. Hence, the expected payoff for player 6 from state t_i equals

$$1 - q_i + q_i \cdot \frac{i-1}{i} = 1 - \frac{q_i}{i} = \frac{i + 1 - q_i \cdot \frac{i+1}{i}}{i+1} .$$

If $q_i > i/(i+1)$, then this number would be strictly smaller than $i/(i+1)$, and player 6 would be better off by leaving the game at s_i. On the other hand, the expected payoff for player 7 from state t_i equals $q_i(n+1)/i$. If $q_i < i/(i+1)$, then this number would be strictly smaller than $(n+1)/(i+1)$, and player 7 would prefer to leave the game at r_i. In both cases, we have a contradiction to $\overline{\sigma}$ being a Nash equilibrium. Hence, $q_i = i/(i+1)$ for all $i = 1, \ldots, n$, and the probability of reaching the game $\mathcal{G}(p_i)$ from s_n equals

$$(1 - q_i) \prod_{j=i+1}^{n} q_i = (1 - \frac{i}{i+1}) \prod_{j=i+1}^{n} \frac{j}{j+1} = \frac{1}{i+1} \cdot \frac{i+1}{n+1} = \frac{1}{n+1} .$$

It remains to be shown that $(\mathcal{G}, s_n)$ has a stationary Nash equilibrium with payoff $\geq (0, \frac{k}{d(n+1)}, 0, \ldots, 0)$ if and only if $\sum_{i=1}^{n} \sqrt{d_i} \geq k$. By Lemma 12, the maximal payoff player 1 receives in a stationary Nash equilibrium of $(\mathcal{G}(p_i), s_1)$ where player 0 receives payoff at least 0 equals $\sqrt{p_i} = \sqrt{d_i}/d$. Hence, the maximal payoff player 1 receives in a stationary Nash equilibrium of $(\mathcal{G}, s_n)$ where player 0 receives payoff at least 0 equals

$$\sum_{i=1}^{n} \frac{1}{n+1} \cdot \frac{\sqrt{d_i}}{d} = \frac{1}{d(n+1)} \cdot \sum_{i=1}^{n} \sqrt{d_i} .$$

We conclude that $(\mathcal{G}, s_n)$ has a stationary Nash equilibrium with payoff $\geq (0, \frac{k}{d(n+1)}, 0, \ldots, 0)$ if and only if $\sum_{i=1}^{n} \sqrt{d_i} \geq k$. $\square$

Again, we can combine our reduction with the game from Example 1 to prove a stronger result for games that are not turn-based.

Corollary 13. Deciding whether a concurrent 8-player terminal reward game has a stationary Nash equilibrium is hard for SqrtSum.

Proof. The proof is analogous to the proof of Corollary 7, but we use the game $\mathcal{G}_1$ from Example 1 instead of the game $\mathcal{G}_2$, and player 0 receives reward 0 in each state of $\mathcal{G}_1$ and reward -1 in the new terminal state. Since $\mathcal{G}_1$ is a terminal-reward game, the resulting game $\mathcal{G}'$ is a terminal-reward game if the original game $\mathcal{G}$ is a terminal-reward game. □

Remark 14. The positive results of Sections 4 and 5 can easily be extended to equilibria in pure or randomised strategies with a memory of a fixed size $k \in \mathbb{N}$: a nondeterministic algorithm can guess a memory structure $\mathfrak{M}$ of size k and then look for a positional, respectively stationary, equilibrium in the product of the original game $\mathcal{G}$ with the memory $\mathfrak{M}$. Hence, for any fixed $k \in \mathbb{N}$, we can decide in PSPACE (NP) the existence of a randomised (pure) equilibrium of size k with payoff $\geq \bar{x}$ and $\leq \bar{y}$. Moreover, these results extend to stochastic games (by appealing to results on MDPs with limit-average objectives; see e.g. [22]).

6 Pure Strategies

In this section, we show that PureNE is decidable and, in fact, NP-complete. Let $\mathcal{G}$ be a concurrent game, $s \in S$ and $i \in \Pi$. We define

$$\mathrm{pval}_i^{\mathcal{G}}(s) = \inf_{\bar{\sigma}} \sup_{\tau} \mathrm{E}_s^{\bar{\sigma}_{-i},\tau}(\phi_i),$$

where $\bar{\sigma}$ ranges over all *pure* strategy profiles of $\mathcal{G}$ and τ ranges over all strategies of player i. Intuitively, $\mathrm{pval}_i^{\mathcal{G}}(s)$ is the lowest payoff that the coalition $\Pi \setminus \{i\}$ can inflict on player i by playing a pure strategy.

By a reduction to a turn-based two-player zero-sum game, we can show that there is a positional strategy profile that attains this value.

Proposition 15. Let $\mathcal{G}$ be a concurrent game, and $i \in \Pi$. There exists a positional strategy profile $\bar{\sigma}^*$ such that $\mathrm{E}_s^{\bar{\sigma}^*_{-i},\tau}(\phi_i) \leq \mathrm{pval}_i^{\mathcal{G}}(s)$ for all states s and all strategies τ of player i.

Proof. We define a turn-based two-player zero-sum game $\mathcal{G}'$ with players 0 and 1 as follows: The set of states of $\mathcal{G}'$ is $S' = S \cup (S \times \Gamma^{\Pi})$. At a state $s \in S$, player 1 chooses an action profile $\bar{a}$ that is legal at s, which leads the game to the state $(s, \bar{a})$. At a state of the form $(s, \bar{a})$, player 0 chooses an action $b \in \Gamma_i(s)$, which leads the game to the state $\delta(s, (\bar{a}_{-i}, b))$. Finally, player 0's reward at a state $s \in S$ or $(s, \bar{a}) \in S \times \Gamma^{\Pi}$ is $r'(s) = r'(s, \bar{a}) = r_i(s)$ (and player 1's reward is the opposite). By [10], there exists a function $\nu \colon S' \to \mathbb{Q}$ (the *value* function) and positional strategies σ^* and τ^* for player 1 and player 0, respectively, such that $\mathrm{E}_s^{\tau,\sigma^*}(\phi'_0) \leq \nu(s)$ for all $s \in S'$ and all strategies τ of player 0 in $\mathcal{G}'$, and

$\mathrm{E}_s^{\tau^*,\sigma}(\phi_0') \geq \nu(s)$ for all $s \in S'$ and all strategies σ of player 1 in $\mathcal{G}'$. We can translate player 1's strategy σ^* into a positional strategy profile $\bar{\sigma}^*$ of $\mathcal{G}$ such that $\mathrm{E}_s^{\bar{\sigma}^*_{-i},\tau}(\phi_i) \leq \nu(s)$ for all states $s \in S$ and all strategies τ of player i in $\mathcal{G}$. Hence, $\mathrm{pval}_i^{\mathcal{G}}(s) \leq \sup_\tau \mathrm{E}_s^{\bar{\sigma}^*_{-i},\tau}(\phi_i) \leq \nu(s)$ for all $s \in S$. We claim that $\mathrm{pval}_i^{\mathcal{G}}(s) \geq \nu(s)$ for all $s \in S$, which implies that $\mathrm{pval}_i^{\mathcal{G}}(s) = \nu(s)$ for all $s \in S$ and that $\bar{\sigma}^*$ is the strategy profile we are looking for. Otherwise, there would exist a pure strategy profile $\bar{\sigma}$ in $\mathcal{G}$ such that $\sup_\tau \mathrm{E}_s^{\bar{\sigma}_{-i},\tau}(\phi_i) < \nu(s)$ for some $s \in S$. But we could translate such a strategy profile $\bar{\sigma}$ into a pure strategy σ of player 1 in $\mathcal{G}'$ such that $\mathrm{E}_s^{\tau^*,\sigma}(\phi_0') < \nu(s)$, a contradiction to the optimality of τ^*. ☐

Given a payoff vector $\bar{z} \in (\mathbb{R} \cup \{\pm\infty\})^{\Pi}$, we define a directed graph $G(\bar{z}) = (V, E)$ (with self-loops) as follows: $V = S$, and there is an edge from s to t if and only if there is an action profile $\bar{a}$ with $\delta(s, \bar{a}) = t$ such that (1) $\bar{a}$ is legal at s and (2) $\mathrm{pval}_i^{\mathcal{G}}(\delta(s, (\bar{a}_{-i}, b))) \leq z_i$ for each player i and each action $b \in \Gamma_i(s)$. Following [3], we call any $\bar{a}$ that fulfils (1) and (2) *$\bar{z}$-secure* at s.

Lemma 16. Let $\bar{z} \in (\mathbb{R} \cup \{\pm\infty\})^{\Pi}$. If there exists an infinite path π in $G(\bar{z})$ from s_0 with $z_i \leq \phi_i(\pi)$ for each player i, then $(\mathcal{G}, s_0)$ has a pure Nash equilibrium with payoff $\phi_i(\pi)$ for player i.

Proof. Let $\pi = s_0 s_1 \ldots$ be an infinite path in $G(\bar{z})$ from s_0 with $z_i \leq \phi_i(\pi)$ for each player i. We define a pure strategy profile $\bar{\sigma}$ as follows: For histories of the form $x = s_0 \bar{a}_0 s_1 \ldots s_{k-1} \bar{a}_{k-1} s_k$, we set $\bar{\sigma}(x)$ to an action profile $\bar{a}$ with $\delta(s_k, \bar{a}) = s_{k+1}$ that is $\bar{z}$-secure at s_k. For all other histories $x = t_0 \bar{a}_0 t_1 \ldots t_{k-1} \bar{a}_{k-1} t_k$, consider the least j such that $s_{j+1} \neq t_{j+1}$. If $\bar{a}_j$ differs from a $\bar{z}$-secure action profile $\bar{a}$ at s_j in precisely one entry i, we set $\bar{\sigma}(x) = \bar{\sigma}^*(t_k)$, where $\bar{\sigma}^*$ is a (fixed) positional strategy profile such that $\mathrm{E}_s^{\bar{\sigma}^*_{-i},\tau}(\phi_i) \leq \mathrm{pval}_i^{\mathcal{G}}(s)$ for all $s \in S$ (which is guaranteed to exist by Proposition 15); otherwise, $\bar{\sigma}(x)$ can be chosen arbitrarily. It is easy to see that $\bar{\sigma}$ is a Nash equilibrium with induced play π. ☐

Lemma 17. Let $\bar{\sigma}$ be a pure Nash equilibrium of $(\mathcal{G}, s_0)$ with payoff $\bar{z}$. Then there exists an infinite path π in $G(\bar{z})$ from s_0 with $\phi_i(\pi) = z_i$ for each player i.

Proof. Let $s_0 \bar{a}_0 s_1 \bar{a}_1 \ldots$ be the play induced by $\bar{\sigma}$. We claim that $\pi := s_0 s_1 \ldots$ is a path in $G(\bar{z})$. Otherwise, consider the least k such that (s_k, s_{k+1}) is not an edge in $G(\bar{z})$. Hence, there exists no $\bar{z}$-secure action profile at $s := s_k$. Since $\bar{a}_k$ is certainly legal at s, there exists a player i and an action $b \in \Gamma_i(s)$ such that $\mathrm{pval}_i^{\mathcal{G}}(\delta(s, (\bar{a}_{-i}, b))) > z_i$. But then player i can improve her payoff by switching to a strategy that mimics σ_i until s is reached, then plays action b, and after that mimics a strategy that ensures payoff $> z_i$ against any pure strategy profile. This contradicts the assumption that $\bar{\sigma}$ is a Nash equilibrium. ☐

Using Lemmas 16 and 17, we can reduce the task of finding a pure Nash equilibrium to the task of finding a path in a multi-weighted graph whose

limit-average weight vector falls between two thresholds. The latter problem can be solved in polynomial time by solving a linear programme with one variable for each pair of a weight function and an edge in the graph, as we prove in the appendix.

Theorem 18. Given a finite directed graph $G = (V, E)$ with weight functions $r_0, \ldots, r_{k-1} \colon V \to \mathbb{Q}$, $v_0 \in V$, and $\bar{x}, \bar{y} \in (\mathbb{Q} \cup \{\pm\infty\})^k$, we can decide in polynomial time whether there exists an infinite path $\pi = v_0 v_1 \ldots$ in G with $x_i \leq \liminf_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} r_i(v_j) \leq y_i$ for all $i = 0, \ldots, k-1$.

We can now describe a nondeterministic algorithm to decide the existence of a pure Nash equilibrium with payoff $\geq \bar{x}$ and $\leq \bar{y}$ in polynomial time. The algorithm starts by guessing, for each player i, a positional strategy profile $\bar{\sigma}^i$ of $\mathcal{G}$ and computes $p_i(s) := \sup_\tau \mathrm{E}_s^{\bar{\sigma}^i_{-i},\tau}(\phi_i)$ for each $s \in S$; these numbers can be computed in polynomial time using the algorithm given by Karp [18]. The algorithm then guesses a vector $\bar{z} \in (\mathbb{R} \cup \{\pm\infty\})^\Pi$ by setting z_i either to x_i or to $p_i(s)$ for some $s \in S$ with $x_i \leq p_i(s)$, and constructs the graph $G'(\bar{z})$, which is defined as $G(\bar{z})$ but with $p_i(s)$ substituted for $\mathrm{pval}_i^{\mathcal{G}}(s)$. Finally, the algorithm determines (in polynomial time) whether there exists an infinite path π in $G(\bar{z})$ from s_0 with $z_i \leq \phi_i(\pi) \leq y_i$ for all $i \in \Pi$. If such a path exists, the algorithm accepts; otherwise it rejects.

Theorem 19. PureNE is in NP.

Proof. We claim that the algorithm described above is correct, i.e. sound and complete. To prove soundness, assume that the algorithm accepts its input. Hence, there exists an infinite path π in $G'(\bar{z})$ from s_0 with $z_i \leq \phi_i(\pi) \leq y_i$. Since $\mathrm{pval}_i^{\mathcal{G}}(s) \leq p_i(s)$ for all $i \in \Pi$ and $s \in S$, the graph $G'(\bar{z})$ is a subgraph of $G(\bar{z})$. Hence, π is also an infinite path in $G(\bar{z})$. By Lemma 16, we can conclude that $(\mathcal{G}, s_0)$ has a pure Nash equilibrium with payoff $\geq \bar{z} \geq \bar{x}$ and $\leq \bar{y}$.

To prove that the algorithm is complete, let $\bar{\sigma}$ be a pure Nash equilibrium of $(\mathcal{G}, s_0)$ with payoff $\bar{z}$, where $\bar{x} \leq \bar{z} \leq \bar{y}$. By Proposition 15, the algorithm can guess positional strategy profiles $\bar{\sigma}^i$ such that $p_i(s) = \mathrm{pval}_i^{\mathcal{G}}(s)$ for all $s \in S$. If the algorithm additionally guesses the payoff vector $\bar{z}'$ defined by $z_i' = \max\{x_i, \mathrm{pval}_i^{\mathcal{G}}(s) : s \in S, \mathrm{pval}_i^{\mathcal{G}}(s) \leq z_i\}$ for all $i \in \Pi$, then the graph $G(\bar{z})$ coincides with the graph $G(\bar{z}')$ (and thus with $G'(\bar{z}')$). By Lemma 17, there exists an infinite path π in $G(\bar{z})$ from s_0 such that $z_i' \leq z_i = \phi_i(\pi) \leq y_i$ for all $i \in \Pi$. Hence, the algorithm accepts. □

The following theorem shows that PureNE is NP-hard. In fact, NP-hardness holds even for turn-based games with rewards 0 and 1.

Theorem 20. PureNE is NP-hard, even for turn-based games with rewards 0 and 1.

Proof. Again, we reduce from SAT. Given a Boolean formula $\varphi = C_1 \wedge \cdots \wedge C_m$ in conjunctive normal form over propositional variables $X_1, \ldots, X_n$, where w.l.o.g. $m \geq 1$ and each clause is nonempty, let $\mathcal{G}$ be the turn-based game described in the proof of Theorem 9 and depicted in Figure 6. We claim that the following statements are equivalent:

1. φ is satisfiable.
2. $(\mathcal{G}, C_1)$ has a positional Nash equilibrium with payoff ≥ 1 for player 0.
3. $(\mathcal{G}, C_1)$ has a pure Nash equilibrium with payoff ≥ 1 for player 0.

Since the implication (1. $\Rightarrow$ 2.) was already proved in the proof of Theorem 9 and the implication (2. $\Rightarrow$ 3.) is trivial, we only need to prove that 3. implies 1. Hence, assume that $(\mathcal{G}, C_1)$ has a pure Nash equilibrium $\overline{\sigma}$ with payoff ≥ 1 for player 0. Since player 0 receives payoff ≥ 1, the terminal state $\perp$ is not reached in the induced play π. Consider the variable assignment α that maps X_i to true if and only if player i receives payoff < 1 from π; we claim that α satisfies the formula. Consider any clause C. Set $T = \{(C, X_i), (C, \neg X_i) : i = 1, \ldots, n\}$, and denote by $\mathbb{1}_s$ the characteristic function of $s \in T$. We have

$$\sum_{s \in T} \liminf_{n \to \infty} \frac{1}{n} \sum_{j=0}^{n-1} -\mathbb{1}_s(\pi(j)) \leq \liminf_{n \to \infty} \frac{1}{n} \sum_{j=0}^{n-1} \sum_{s \in T} -\mathbb{1}_s(\pi(j)) = -\frac{1}{2m} < 0\,.$$

In particular, there exists a state $s = (C, L)$ such that

$$\liminf_{n \to \infty} \frac{1}{n} \sum_{j=0}^{n-1} -\mathbb{1}_s(\pi(j)) < 0\,.$$

If $L = X_i$, then $r_i \leq 1 - \mathbb{1}_s$. Hence, $\phi_i(\pi) < 1$, and α maps X_i to true, thereby satisfying C. If $L = \neg X_i$, then player i must receive payoff 1, because otherwise she could improve her payoff by playing from s to $\perp$. Hence, α maps X_i to false and satisfies C. $\square$

It follows from Theorems 19 and 20 that PureNE is NP-complete. By combining our reduction with a game that has no pure Nash equilibrium, we can prove the following stronger result for non-turn-based games.

Corollary 21. Deciding the existence of a pure Nash equilibrium in a concurrent limit-average game is NP-complete, even for games with rewards 0 and 1.

Proof. The proof is analogous to the proof of Corollary 10. $\square$

Note that Theorem 20 and Corollary 21 do not apply to terminal-reward games. In fact, PureNE is decidable in P for these games, which follows from two facts about terminal-reward games: (1) the numbers $\mathrm{pval}_i^{\mathcal{G}}(s)$ can be computed in polynomial time (using a reduction to a turn-based two-player zero-sum game and applying a result of Washburn [31]), and (2) the only possible vectors that

can emerge as the payoff of a pure strategy profile are the zero vector and the reward vectors at terminal states.

Theorem 22. PureNE is in P for terminal-reward games.

7 Randomised Strategies

In this section, we show that the problem NE is undecidable and, in fact, not recursively enumerable for turn-based terminal-reward games. The proof proceeds by a reduction from an undecidable problem about *two-counter machines*. Such a machine is of the form $\mathcal{M} = (Q, q_0, \Delta)$, where

- Q is a finite set of *states*,
- $q_0 \in Q$ is the *initial state*,
- $\Delta \subseteq Q \times \Gamma \times Q$ is a set of *transitions*.

The set Γ specifies which instructions $\mathcal{M}$ may perform on its counters. For our purposes, the instruction set $\Gamma := \{\mathrm{inc}(j), \mathrm{dec}(j), \mathrm{zero}(j) : j = 1, 2\}$ suffices: a counter can be incremented, decremented, or tested for zero. For $q \in Q$ we write $q\Delta$ for the set of all $(\gamma, q') \in \Gamma \times Q$ such that $(q, \gamma, q') \in \Delta$. The machine $\mathcal{M}$ is *deterministic* if for each $q \in Q$ either (1) $q\Delta = \emptyset$, (2) $q\Delta = \{(\mathrm{inc}(j), q')\}$ for some $j \in \{1, 2\}$ and $q' \in Q$, or (3) $q\Delta = \{(\mathrm{zero}(j), q_1), (\mathrm{dec}(j), q_2)\}$ for some $j \in \{1, 2\}$ and $q_1, q_2 \in Q$.

A configuration of $\mathcal{M}$ is a triple $C = (q, i_1, i_2) \in Q \times \mathbb{N} \times \mathbb{N}$, where q denotes the current state and i_j denotes the current value of counter j. A configuration $C' = (q', i_1', i_2')$ is a *successor* of configuration $C = (q, i_1, i_2)$, denoted by $C \vdash C'$, if there exists a "matching" transition $(q, \gamma, q') \in \Delta$. For example, $(q, i_1, i_2) \vdash (q', i_1 + 1, i_2)$ if and only if $(q, \mathrm{inc}(1), q') \in \Delta$. The instruction $\mathrm{zero}(j)$ performs a *zero test*: $(q, i_1, i_2) \vdash (q', i_1, i_2)$ if and only if $i_1 = 0$ and $(q, \mathrm{zero}(1), q') \in \Delta$, or $i_2 = 0$ and $(q, \mathrm{zero}(2), q') \in \Delta$.

A *partial computation of* $\mathcal{M}$ is a sequence $\rho = \rho(0)\rho(1)\dots$ of configurations such that $\rho(0) \vdash \rho(1) \vdash \cdots$ and $\rho(0) = (q_0, 0, 0)$ (the *initial configuration*). A partial computation of $\mathcal{M}$ is a *computation of* $\mathcal{M}$ if it is infinite or it ends in a configuration C for which there is no C' with $C \vdash C'$. Note that each deterministic two-counter machine has a unique computation.

The *halting problem* is to decide, given a machine $\mathcal{M}$, whether the computation of $\mathcal{M}$ is finite. It is well-known that deterministic two-counter machines are Turing powerful, which makes the halting problem and its dual, the *non-halting problem*, undecidable, even when restricted to deterministic two-counter machines. In fact, the non-halting problem for deterministic two-counter machines is not recursively enumerable.

To prove the undecidability of NE, we employ a reduction from the non-halting problem for deterministic two-counter machines. More precisely, we show how to compute from such a machine $\mathcal{M}$ a game $(\mathcal{G}, s_0)$ such that the

computation of $\mathcal{M}$ is infinite if and only if there exists a Nash equilibrium of $(\mathcal{G}, s_0)$ where player 0 receives expected payoff ≥ 0. Without loss of generality, we assume that in $\mathcal{M}$ there is no zero test that is followed by another zero test: if $(q, \mathrm{zero}(j), q') \in \Delta$, then $|q'\Delta| \leq 1$.

The game $\mathcal{G}$ is played by players 0, 1 and 12 other players A_j^t, B_j^t, D^t and E_j, indexed by $j \in \{1,2\}$ and $t \in \{0,1\}$. Intuitively, player 0 and player 1 build up the computation of $\mathcal{M}$: player 0 updates the counters, and player 1 chooses transitions. Players A_j^t and B_j^t make sure that player 0 updates the counters correctly: players A_j^0 and A_j^1 ensure that, in each step, the value of counter j is not too high, and players B_j^0 and B_j^1 ensure that, in each step, the value of counter j is not too low. More precisely, A_j^0 and B_j^0 monitor the even steps of the computation, while A_j^1 and B_j^1 monitor the odd steps. Finally, players D^t and E_j ensure that player 0 uses a randomised strategy of a restricted form.

Let $\Gamma' := \Gamma \cup \{\mathrm{init}\}$. For each $q \in Q$, each $\gamma \in \Gamma'$, each $j \in \{1,2\}$ and each $t \in \{0,1\}$, the game $\mathcal{G}$ contains the gadgets $S_{\gamma,q}^t$, I_q^t and $C_{\gamma,j}^t$, which are depicted in Figure 9. The initial state of $\mathcal{G}$ is $s_0 := s_{\mathrm{init},q_0}^0$. Note that in the gadget $S_{\gamma,q}^t$, each of the players A_j^t, B_j^t, D^t and E_j may unilaterally decide to *quit the game*, which gives the respective player a payoff of 1 or 2, but payoff -1 to player 0.

It will turn out that player 1 will play a pure strategy in any Nash equilibrium of $(\mathcal{G}, s_0)$ where player 0 receives expected payoff 0, except possibly for histories that are not consistent with the equilibrium. Moreover, player 0 has to play a uniform distribution inside $S_{\gamma,q}^t$. Formally, we say that a strategy profile $\bar{\sigma}$ of $\mathcal{G}$ is *safe* if 1. $\sigma_0(xs)$ assigns probability $\frac{1}{2}$ to both outgoing transitions for all histories xs consistent with $\bar{\sigma}$ and ending in a state $s \in S_{\gamma,q}^t$ controlled by player 0, and 2. $\sigma_1(xs)$ is degenerate for all histories xs consistent with $\bar{\sigma}$ and ending in a state s controlled by player 1.

For each safe strategy profile $\bar{\sigma}$ where player 0 receives expected payoff 0, let $x_0 s_0 \prec x_1 s_1 \prec x_2 s_2 \prec \ldots$ ($x_i \in S^*$, $s_i \in S$, $x_0 = \varepsilon$) be the unique sequence consisting of all histories xs of $(\mathcal{G}, s_0)$ consistent with $\bar{\sigma}$ that end in a state s of the form $s = s_{\gamma,q}^t$. This sequence is infinite because $\bar{\sigma}$ is safe and player 0 receives expected payoff 0. Additionally, let $q_0, q_1, \ldots$ be the corresponding sequence of states and $\gamma_0, \gamma_1, \ldots$ be the corresponding sequence of instructions, i.e. $s_n = s_{\gamma_n,q_n}^0$ or $s_n = s_{\gamma_n,q_n}^1$ for all $n \in \mathbb{N}$. For each $j \in \{1,2\}$ and $n \in \mathbb{N}$, we define two conditional expectations as follows:

$$a_j^n := \mathrm{E}_{s_0}^{\bar{\sigma}}(\phi_{A_j^{n \bmod 2}} \mid x_n s_n \cdot S^\omega)\,;$$
$$b_j^n := \mathrm{E}_{s_0}^{\bar{\sigma}}(\phi_{B_j^{n \bmod 2}} \mid x_n s_n \cdot S^\omega)\,.$$

Note that at every terminal state of the counter gadgets $C_{\gamma,j}^t$ and $C_{\gamma,j}^{1-t}$ the rewards of player A_j^t and player B_j^t sum up to 4. For each j, the conditional probability that, given the history $x_n s_n$, we reach such a state is $\sum_{k \in \mathbb{N}} \frac{1}{2^k} \cdot \frac{1}{4} = \frac{1}{2}$. Hence,

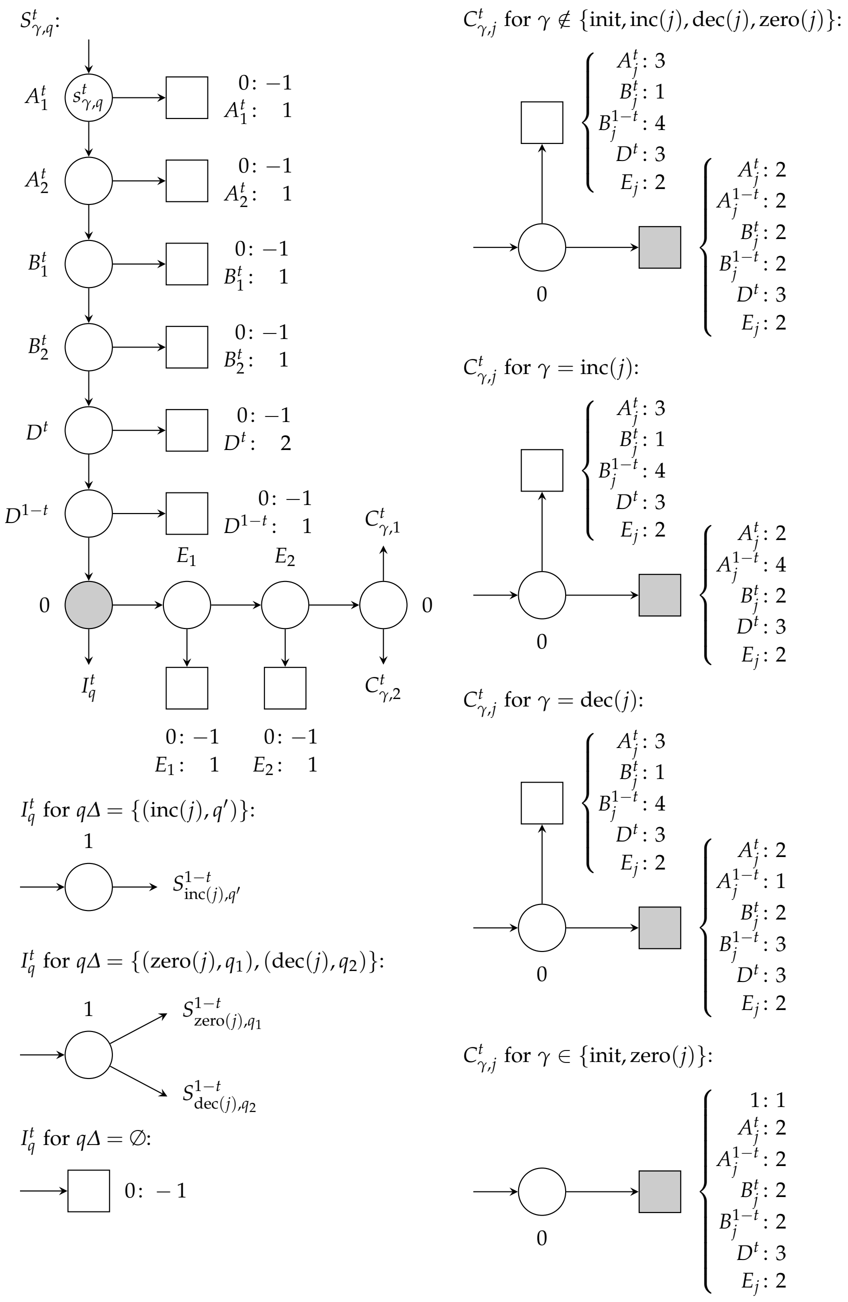

Figure 9. Simulating a two-counter machine.

$a_j^n + b_j^n = 2$ for all $n \in \mathbb{N}$. We say that $\bar{\sigma}$ is *stable* if $a_j^n = 1$ or, equivalently, $b_j^n = 1$ for each $j \in \{1,2\}$ and for all $n \in \mathbb{N}$.

Finally, for each $j \in \{1,2\}$ and $n \in \mathbb{N}$, we define a number $c_j^n \in [0,1]$ as follows: After the history $x_n s_n$, with probability $\frac{1}{4}$ the play proceeds to the state controlled by player 0 in the counter gadget $C_{\gamma_n,j}^{n \bmod 2}$. The number c_j^n is defined as the probability that player 0 plays to the neighbouring grey state. Note that, by the construction of $\mathcal{G}$, it holds that $c_j^n = 1$ if $\gamma_n = \mathrm{zero}(j)$ or $\gamma_n = \mathrm{init}$. In particular, $c_1^0 = c_2^0 = 1$.

Lemma 23. Let $\bar{\sigma}$ be a safe strategy profile with expected payoff 0 for player 0. Then $\bar{\sigma}$ is stable if and only if

$$c_j^{n+1} = \begin{cases} \frac{1}{2} \cdot c_j^n & \text{if } \gamma_{n+1} = \mathrm{inc}(j), \\ 2 \cdot c_j^n & \text{if } \gamma_{n+1} = \mathrm{dec}(j), \\ c_j^n = 1 & \text{if } \gamma_{n+1} = \mathrm{zero}(j), \\ c_j^n & \text{otherwise.} \end{cases} \tag{1}$$

for each $j \in \{1,2\}$ and for all $n \in \mathbb{N}$.

To prove the lemma, consider a safe strategy profile $\bar{\sigma}$ of $\mathcal{G}$ with expected payoff 0 for player 0. For each $j \in \{1,2\}$ and $n \in \mathbb{N}$, we define yet another conditional expectation

$$p_j^n := \mathrm{E}_{s_0}^{\bar{\sigma}}(\phi_{A_j^{n \bmod 2}} \mid x_n s_n \cdot S^\omega \setminus x_{n+2} s_{n+2} \cdot S^\omega)\,.$$

The following claim relates the numbers a_j^n and p_j^n.

Claim. Let $j \in \{1,2\}$. Then $a_j^n = 1$ for all $n \in \mathbb{N}$ if and only if $p_j^n = \frac{3}{4}$ for all $n \in \mathbb{N}$.

Proof. ($\Rightarrow$) Assume that $a_j^n = 1$ for all $n \in \mathbb{N}$. We have $a_j^n = p_j^n + \frac{1}{4} \cdot a_j^{n+2}$ and therefore $1 = p_j^n + \frac{1}{4}$ for all $n \in \mathbb{N}$. Hence, $p_j^n = \frac{3}{4}$ for all $n \in \mathbb{N}$.

($\Leftarrow$) Assume that $p_j^n = \frac{3}{4}$ for all $n \in \mathbb{N}$. Since $a_j^n = p_j^n + \frac{1}{4} \cdot a_j^{n+2}$ for all $n \in \mathbb{N}$, the numbers a_j^n have to satisfy the following recurrence: $a_j^{n+2} = 4a_j^n - 3$. Since all the numbers a_j^n are bounded by the minimum and maximum reward for player $A_j^{n \bmod 2}$, we have $0 \leq a_j^n \leq 4$ for all $n \in \mathbb{N}$. It is easy to see that the only values for a_j^0 and a_j^1 such that $0 \leq a_j^n \leq 4$ for all $n \in \mathbb{N}$ are $a_j^0 = a_j^1 = 1$. But this implies that $a_j^n = 1$ for all $n \in \mathbb{N}$. □

Proof (of Lemma 23). By the previous claim, it suffices to show that $p_j^n = \frac{3}{4}$ if and only if (1) holds. Let $j \in \{1,2\}$, $n \in \mathbb{N}$ and $t = n \bmod 2$. The number p_j^n can be expressed as a weighted average of the expected payoff for player A_j^t inside $C_{\gamma_n,j}^t$ and the expected payoff for player A_j^t inside $C_{\gamma_{n+1},j}^{1-t}$. The first payoff does not

depend on γ_n, but the second depends on γ_{n+1}. Let us consider the case that $\gamma_{n+1} = \mathrm{inc}(j)$. In this case, p_j^n equals

$$\tfrac{1}{4} \cdot \big(c_j^n \cdot 2 + (1 - c_j^n) \cdot 3\big) + \tfrac{1}{8} \cdot c_j^{n+1} \cdot 4 = \tfrac{3}{4} - \tfrac{1}{4} \cdot c_j^n + \tfrac{1}{2} \cdot c_j^{n+1}.$$

Obviously, this sum equals $\frac{3}{4}$ if and only if $c_j^{n+1} = \frac{1}{2} \cdot c_j^n$. For any other value of γ_{n+1}, the argumentation is similar. □

The next lemma states that every Nash equilibrium with expected payoff 0 for player 0 is, in fact, safe.

Lemma 24. Let $\bar{\sigma}$ be a Nash equilibrium of $(\mathcal{G}, s_0)$ with expected payoff 0 for player 0. Then $\bar{\sigma}$ is safe.

Proof. We start by proving that player 0 plays a uniform distribution inside $S^t_{\gamma,q}$. We prove this separately for histories that end in a white state and histories that end in a grey state.

Let xs be a history consistent with $\bar{\sigma}$ and ending in a white state $s \in S^t_{\gamma,q}$ controlled by player 0. Since the players E_1 and E_2 can ensure payoff 1 by quitting the game, player 0 has to play to $C^t_{\gamma,1}$ and $C^t_{\gamma,2}$ with probability $\frac{1}{2}$ each. Otherwise, $\bar{\sigma}$ would not be a Nash equilibrium.

Now let xs be a history consistent with $\bar{\sigma}$ and ending in a grey state $s \in S^t_{\gamma,q}$ controlled by player 0. In the following, let $t = 0$; the proof for $t = 1$ is analogous. Denote by p the probability that player 0 plays to $t \in I^t_q$ after the history xs. For $i \in \{0, 1\}$, let

$$d^i = \mathrm{E}^{\bar{\sigma}}_{s_0}(\phi_{D^i} \mid xst \cdot S^\omega).$$

By the definition of the game, we have $d^0 \geq 1$ and $d^1 \geq 2$. On the other hand, since at every terminal state the sum of the rewards for players D^0 and D^1 is at most 3, we have $d^0 + d^1 \leq 3$. Hence, $d^0 = 1$ and $d^1 = 2$. Consider the expected payoffs for players D^0 and D^1 after the history xs:

$$\begin{aligned}
&\mathrm{E}^{\bar{\sigma}}_{s_0}(\phi_{D^0} \mid xs \cdot S^\omega) = (1-p) \cdot 3 + p \cdot d^0 = 3 - 2p;\\
&\mathrm{E}^{\bar{\sigma}}_{s_0}(\phi_{D^1} \mid xs \cdot S^\omega) = p \cdot d^1 = 2p\,.
\end{aligned}$$

Since $\bar{\sigma}$ is a Nash equilibrium, these numbers are bounded from below by 2 and 1, respectively (otherwise, it would be better for player D^0 or D^1 to quit the game). Hence, $p = \frac{1}{2}$.

To prove that $\bar{\sigma}$ is safe, it remains to be shown that player 1 plays a degenerate distribution for all histories xs consistent with $\bar{\sigma}$ and ending in a state $s \in I^t_q$. Towards a contradiction, assume that xs is such a history and that $\sigma_1(xs)$ assigns probability > 0 to two distinct successor states. Hence, $q\Delta = \{(\mathrm{zero}(j), q_1), (\mathrm{dec}(j), q_2)\}$ for some $j \in \{1, 2\}$ and $q_1, q_2 \in Q$. By our

assumption that there are no consecutive zero tests and since player 0 receives expected payoff 0,

$$\mathrm{E}^{\bar{\sigma}}_{s_0}(\phi_1 \mid xs \cdot s^{1-t}_{\mathrm{zero}(j),q_1} \cdot S^{\omega}) \geq \tfrac{1}{4}\,,$$

but

$$\mathrm{E}^{\bar{\sigma}}_{s_0}(\phi_1 \mid xs \cdot s^{1-t}_{\mathrm{dec}(j),q_2} \cdot S^{\omega}) \leq \tfrac{1}{6}\,.$$

Hence, player 1 could improve her payoff by playing to $s^{1-t}_{\mathrm{zero}(j),q_1}$ with probability 1, a contradiction to $\bar{\sigma}$ being a Nash equilibrium. ☐

Finally, we can prove the following theorem.

Theorem 25. NE is not recursively enumerable, even for turn-based 14-player terminal-reward games.

Proof. We claim that the function mapping a deterministic two-counter machine $\mathcal{M}$ to the 14-player game $(\mathcal{G}, s_0)$ as described above realises a many-one reduction from the non-halting problem to NE. Clearly, $\mathcal{G}$ can be computed from $\mathcal{M}$. We prove that the computation of $\mathcal{M}$ is infinite if and only if $(\mathcal{G}, s_0)$ has a Nash equilibrium in which player 0 receives expected payoff (at least) 0.

($\Rightarrow$) Assume that the computation $\rho = \rho(0)\rho(1)\ldots$ of $\mathcal{M}$ is infinite. Player 0's equilibrium strategy σ_0 can be described as follows: For a history that ends at the unique state controlled by player 0 in the gadget $C^t_{\gamma,j}$ after visiting a state of the form $s^t_{\gamma',q}$ or $s^{1-t}_{\gamma',q}$ exactly $n > 0$ times, player 0 plays to the grey successor state with probability 2^{-i}, where i is the value of counter j in configuration $\rho(n-1)$. Moreover, for a history that ends at a state controlled by player 0 in the gadget $S^t_{\gamma,q}$, player 0 plays to both successors with probability $\frac{1}{2}$ each.

The only place where player 1 has a choice is the sole state in the gadget I^t_q for $q\Delta = \{(\mathrm{zero}(j), q_1), (\mathrm{dec}(j), q_2)\}$. If the play arrives at such a state after visiting a state of the form $s^t_{\gamma,q'}$ or $s^{1-t}_{\gamma,q'}$ exactly $n > 0$ times, then player 1's pure strategy σ_1 prescribes to play to $S^{1-t}_{\mathrm{zero}(j),q_1}$ if the value of counter j in configuration $\rho(n-1)$ is zero and to $S^{1-t}_{\mathrm{dec}(j),q_2}$ if the value of counter j in configuration $\rho(n-1)$ is non-zero.

Any other player's pure strategy is defined as follows: After a history ending in $S^t_{\gamma,q}$, the strategy prescribes to quit the game if and only if the history is not compatible with ρ (i.e. the corresponding sequence of instructions does not match ρ).

Note that the resulting strategy profile $\bar{\sigma}$ is safe. Moreover, since player 0 and player 1 follow the computation of $\mathcal{M}$, a terminal state inside one of the counter gadgets $C^t_{\gamma,j}$ is reached with probability 1. Since player 0 receives reward 0 at any such terminal state, player 0's expected payoff equals 0. Finally, by the definition of $\bar{\sigma}$, for each $j \in \{1, 2\}$ and for all $n \in \mathbb{N}$, if i and i' are the values of counter j

in configuration $\rho(n)$ and configuration $\rho(n+1)$, respectively, then $c_j^n = 2^{-i}$, $c_j^{n+1} = 2^{-i'}$, and γ_{n+1} is the instruction corresponding to the counter update from $\rho(n)$ to $\rho(n+1)$. Hence, (1) holds, and we can conclude from Lemma 23 that $\bar{\sigma}$ is stable.

We claim that $\bar{\sigma}$ is, in fact, a Nash equilibrium of $(\mathcal{G}, s_0)$: It is obvious that player 0 cannot improve her payoff. If player 1 deviates, then with positive probability we reach a history that is not compatible with ρ; hence, player A_1^0 or A_1^1 will quit the game, which ensures that player 1 will receive payoff 0 after this history. Since $\bar{\sigma}$ is stable, none of the players A_j^t or B_j^t can improve her payoff. Finally, the expected payoffs of player D^t and player D^{1-t} from $s_{\gamma,q}^t$ equal 2 and 1, respectively, which is the same as they would get if they quit the game. The reasoning for players E_1 and E_2 is analogous.

($\Leftarrow$) Assume that $\bar{\sigma}$ is a Nash equilibrium of $(\mathcal{G}, s_0)$ with expected payoff ≥ 0 for player 0. Since 0 is the maximum reward for player 0, this means that the expected payoff of $\bar{\sigma}$ for player 0 equals 0. From Lemma 24, we can conclude that $\bar{\sigma}$ is safe. To apply Lemma 23 and obtain (1), it remains to be shown that $\bar{\sigma}$ is stable. In order to derive a contradiction, assume that there exists $j \in \{1, 2\}$ and $n \in \mathbb{N}$ such that either $a_j^n < 1$ or $a_j^n > 1$, i.e. $b_j^n < 1$. In the first case, player $A_j^{n \bmod 2}$ could improve her payoff by quitting the game after history $x_n s_n$, while in the second case, player $B_j^{n \bmod 2}$ could improve her payoff by quitting the game, again a contradiction to $\bar{\sigma}$ being a Nash equilibrium.

From (1) and the fact that $c_j^0 = 1$, it follows that each c_j^n is of the form $c_j^n = 2^{-i}$ with $i \in \mathbb{N}$. We denote by i_j^n the unique number i such that $c_j^n = 2^{-i}$ and set $\rho(n) = (q_n, i_1^n, i_2^n)$ for each $n \in \mathbb{N}$. We claim that $\rho := \rho(0)\rho(1)\ldots$ is in fact the computation of $\mathcal{M}$. In particular, this computation is infinite. It suffices to verify the following two properties:

- $\rho(0) = (q_0, 0, 0)$.
- $\rho(n) \vdash \rho(n+1)$ for all $n \in \mathbb{N}$.

The first property is immediate. To prove the second property, let $\rho(n) = (q, i_1, i_2)$ and $\rho(n+1) = (q', i_1', i_2')$. Hence, s_n lies inside $S_{\gamma,q}^t$, and s_{n+1} lies inside $S_{\gamma',q'}^{1-t}$ for suitable γ, γ' and $t = n \bmod 2$. We only prove the claim for $q\Delta = \{(\text{zero}(1), q_1), (\text{dec}(1), q_2)\}$; the other cases are similar. Note that, by the construction of the gadget I_q^t, it must be the case that either $q' = q_1$ and $\gamma' = \text{zero}(1)$, or $q' = q_2$ and $\gamma' = \text{dec}(1)$. By (1), if $\gamma' = \text{zero}(1)$, then $i_1' = i_1 = 0$ and $i_2' = i_2$, and if $\gamma' = \text{dec}(1)$, then $i_1' = i_1 - 1$ and $i_2' = i_2$. This implies $\rho(n) \vdash \rho(n+1)$: On the one hand, if $i_1 = 0$, then $i_1' \neq i_1 - 1$, which implies $\gamma' \neq \text{dec}(1)$ and thus $\gamma' = \text{zero}(1)$, $q' = q_1$ and $i_1' = i_1 = 0$. On the other hand, if $i_1 > 0$, then $\gamma' \neq \text{zero}(1)$ and thus $\gamma' = \text{dec}(1)$, $q' = q_2$ and $i_1' = i_1 - 1$. $\square$

For games that are not turn-based, we can show the stronger theorem that the set of all games that have a Nash equilibrium is not recursively enumerable.

Corollary 26. The set of all initialised concurrent 14-player terminal-reward games that have a Nash equilibrium is not recursively enumerable.

Proof. The proof is analogous to the proof of Corollary 10, but we use the game $\mathcal{G}_1$ from Example 1 instead of the game $\mathcal{G}_2$, and we set the reward for player 0 in each state of $\mathcal{G}_1$ to 0. $\square$

8 Conclusion

We have analysed the complexity of Nash equilibria in concurrent games with limit-average objectives. In particular, we have shown that randomisation in strategies leads to undecidability, while restricting to pure strategies retains decidability. This is in contrast to stochastic games, where pure strategies lead to undecidability [27]. While we have provided matching and lower bounds in most cases, there remain some problems where we do not know the exact complexity. Apart from StatNE, these include the problem PureNE when restricted to a bounded number of players.

References

[1] E. Allender, P. Bürgisser, J. Kjeldgaard-Pedersen & P. B. Miltersen. On the complexity of numerical analysis. *SIAM Journal on Computing*, 38(5):1987–2006, 2009.

[2] R. Alur, A. Degorre, O. Maler & G. Weiss. On omega-languages defined by mean-payoff conditions. In *Proceedings of the 12th International Conference on Foundations of Software Science and Computation Structures, FOSSACS 2009*, volume 5504 of *Lecture Notes in Computer Science*, pages 333–347. Springer-Verlag, 2009.

[3] P. Bouyer, R. Brenguier & N. Markey. Nash equilibria for reachability objectives in multi-player timed games. In *Proceedings of the 21st International Conference on Concurrency Theory, CONCUR 2010*, volume 6269 of *Lecture Notes in Computer Science*, pages 192–206. Springer-Verlag, 2010.

[4] J. Canny. Some algebraic and geometric computations in PSPACE. In *Proceedings of the 20th annual ACM Symposium on Theory of Computing, STOC '88*, pages 460–469. ACM Press, 1988.

[5] K. Chatterjee, L. Doyen, T. A. Henzinger & J.-F. Raskin. Generalized mean-payoff and energy games. In *Proceedings of the 30th International Conference on Foundations of Software Technology and Theoretical Computer Science, FSTTCS 2010*, volume 8 of *Leibniz International Proceedings in Informatics*. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2010.

[6] X. Chen, X. Deng & S.-H. Teng. Settling the complexity of computing two-player Nash equilibria. *Journal of the ACM*, 56(3), 2009.

[7] V. Conitzer & T. Sandholm. Complexity results about Nash equilibria. In *Proceedings of the 18th International Joint Conference on Artificial Intelligence, IJCAI 2003*, pages 765–771. Morgan Kaufmann, 2003.

[8] C. Daskalakis, P. W. Goldberg & C. H. Papadimitriou. The complexity of computing a Nash equilibrium. *SIAM Journal on Computing*, 39(1):195–259, 2009.

[9] L. de Alfaro, T. A. Henzinger & O. Kupferman. Concurrent reachability games. *Theoretical Computer Science*, 386(3):188–217, 2007.

[10] A. Ehrenfeucht & J. Mycielski. Positional strategies for mean payoff games. *International Journal of Game Theory*, 8:109–113, 1979.

[11] K. Etessami & M. Yannakakis. On the complexity of Nash equilibria and other fixed points. *SIAM Journal on Computing*, 39(6):2531–2597, 2010.

[12] H. Everett. Recursive games. In M. Dresher, A. W. Tucker & P. Wolfe (eds.), *Contributions to the Theory of Games III*, volume 39 of *Annals of Mathematical Studies*, pages 47–78. Princeton University Press, 1957.

[13] A. M. Fink. Equilibrium in a stochastic n-person game. *Journal of Science in Hiroshima University*, 28(1):89–93, 1964.

[14] D. Fisman, O. Kupferman & Y. Lustig. Rational synthesis. In *Proceedings of the 16th International Conference on Tools and Algorithms for the Construction and Analysis of Systems, TACAS 2010*, volume 6015 of *Lecture Notes in Computer Science*, pages 190–204. Springer-Verlag, 2010.

[15] M. R. Garey, R. L. Graham & D. S. Johnson. Some NP-complete geometric problems. In *Proceedings of the 8th Annual ACM Symposium on Theory of Computing, STOC '76*, pages 10–22. ACM Press, 1976.

[16] D. Gillette. Stochastic games with zero stop probabilities. In M. Dresher, A. W. Tucker & P. Wolfe (eds.), *Contributions to the Theory of Games III*, volume 39 of *Annals of Mathematical Studies*, pages 179–187. Princeton University Press, 1957.

[17] T. A. Henzinger. Games in system design and verification. In *Proceedings of the 10th Conference on Theoretical Aspects of Rationality and Knowledge, TARK 2005*, pages 1–4. National University of Singapore, 2005.

[18] R. M. Karp. A characterization of the minimum cycle mean in a digraph. *Discrete Mathematics*, 23(3):309–311, 1978.

[19] J.-F. Mertens & A. Neyman. Stochastic games. *International Journal of Game Theory*, 10(2):53–66, 1981.

[20] J. F. Nash, Jr. Equilibrium points in N-person games. *Proceedings of the National Academy of Sciences of the USA*, 36:48–49, 1950.

[21] A. Neyman & S. Sorin (eds.). *Stochastic Games and Applications*, volume 570 of *NATO Science Series C*. Springer-Verlag, 2003.

[22] M. L. Puterman. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. John Wiley and Sons, 1994.

[23] A. Schrijver. *Theory of Linear and Integer Programming*. John Wiley and Sons, 1998.

[24] L. S. Shapley. Stochastic games. *Proceedings of the National Academy of Sciences of the USA*, 39:1095–1100, 1953.

[25] F. Thuijsman & T. E. S. Raghavan. Perfect-information stochastic games and related classes. *International Journal of Game Theory*, 26:403–408, 1997.

[26] M. Ummels. The complexity of Nash equilibria in infinite multiplayer games. In *Proceedings of the 11th International Conference on Foundations of Software Science and Computation Structures, FOSSACS 2008*, volume 4962 of *Lecture Notes in Computer Science*, pages 20–34. Springer-Verlag, 2008.

[27] M. Ummels & D. Wojtczak. The complexity of Nash equilibria in simple stochastic multiplayer games. In *Proceedings of the 36th International Colloquium on Automata, Languages and Programming, ICALP 2009*, volume 5556 of *Lecture Notes in Computer Science*, pages 297–308. Springer-Verlag, 2009.

[28] Y. Velner & A. Rabinovich. Church synthesis problem for noisy input. In *Proceedings of the 14th International Conference on Foundations of Software Science and Computation Structures, FOSSACS 2011*, volume 6604 of *Lecture Notes in Computer Science*, pages 275–289. Springer-Verlag, 2011.

[29] N. Vielle. Two-player stochastic games I: A reduction. *Israel Journal of Mathematics*, 119(1):55–91, 2000.

[30] N. Vielle. Two-player stochastic games II: The case of recursive games. *Israel Journal of Mathematics*, 119(1):93–126, 2000.

[31] A. R. Washburn. Deterministic graphical games. *Journal of Mathematical Analysis and Applications*, 153:84–96, 1990.

[32] U. Zwick & M. Paterson. The complexity of mean payoff games on graphs. *Theoretical Computer Science*, 158(1–2):343–359, 1996.

Appendix

This appendix is devoted to the proof of Theorem 18, which is restated here.

Theorem 18. Given a finite directed graph $G = (V, E)$ with weight functions $r_0, \dots, r_{k-1} \colon V \to \mathbb{Q}$, $v_0 \in V$, and $\bar{x}, \bar{y} \in (\mathbb{Q} \cup \{\pm\infty\})^k$, we can decide in polynomial time whether there exists an infinite path $\pi = v_0 v_1 \dots$ in G with $x_i \leq \liminf_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} r_i(v_j) \leq y_i$ for all $i = 0, \dots, k-1$.

In the following, let $G = (V, E)$ be a finite directed graph with weight functions $r_0, \dots, r_{k-1} \colon V \to \mathbb{Q}$, and set $[k] = \{0, 1, \dots, k-1\}$. Given a vertex $v \in V$, we write $\mathrm{In}(v)$ and $\mathrm{Out}(v)$ for the set of all edges that end, respectively start, in v. Moreover, given an edge $e = (u, v) \in E$ we set $r_i(e) := r_i(u)$. We extend the weight functions r_i to finite paths by setting $r_i(v_1 \dots v_n) = \sum_{j=1}^{n} r_i(v_j)$. If $\pi = \pi(0)\pi(1)\dots$ is an infinite path and $n \in \mathbb{N}$, we write $\pi \upharpoonright n$ for the finite path $\pi(0) \dots \pi(n-1)$, and we set $\phi_i(\pi) := \liminf_{n\to\infty} r_i(\pi \upharpoonright n)/n$, i.e. $\phi_i(\pi)$ is precisely the limit-average weight of the path π w.r.t. the weight function r_i. Finally, $\phi(\pi)$ denotes the vector $(\phi_i(\pi))_{i\in[k]}$. Now consider the following linear constraints over the variables $f_{i,e}$, where $i \in [k]$ and $e \in E$:

(1) $f_{i,e} \geq 0$ for all $i \in [k]$ and $e \in E$;
(2) $\sum_{e\in E} f_{i,e} = 1$ for all $i \in [k]$;
(3) $\sum_{e\in \mathrm{In}(v)} f_{i,e} = \sum_{e\in \mathrm{Out}(v)} f_{i,e}$ for all $i \in [k]$ and $v \in V$;
(4) $x_i \leq \sum_{e\in E} f_{i,e} \cdot r_i(e) \leq y_i$ for all $i \in [k]$;
(5) $\sum_{e\in E} f_{i,e} \cdot r_i(e) \leq \sum_{e\in E} f_{j,e} \cdot r_i(e)$ for all $i, j \in [k]$.

Lemma 27. If there exists an infinite path π in G such that $\bar{x} \leq \phi(\pi) \leq \bar{y}$, then there exists a solution to (1)–(5).

Proof. Let $\pi = \pi(0)\pi(1)\dots$ be an infinite path in G such that $\bar{x} \le \phi(\pi) \le \bar{y}$. Given $n \in \mathbb{N}$ and $e \in E$, define $\kappa(n,e) := |\{j < n : (\pi(j), \pi(j+1)) = e\}|$. Moreover, for $n > 0$, set $\lambda(n,e) = \kappa(n,e)/n$. Note that $0 \le \lambda(n,e) \le 1$ for all $e \in E$ and $n \in \mathbb{N}$. In order to define the numbers $f_{i,e}$, let us now fix $i \in [k]$. Since $\phi_i(\pi) = \liminf_{n\to\infty} r_i(\pi \restriction n)/n$, there exist natural numbers $0 < k_0^i < k_1^i < \cdots$ such that $\phi_i(\pi) = \lim_{n\to\infty} r_i(\pi \restriction k_n^i)/k_n^i$. Now we define a sequence $\varphi_0^i, \varphi_1^i, \dots$ of vectors $\varphi_n^i \in \mathbb{R}^E$ by setting $\varphi_n^i(e) = \lambda(k_n^i, e)$. Since this sequence is bounded, by the Bolzano-Weierstrass theorem, there exists a converging subsequence $\psi_0^i, \psi_1^i, \dots$ of this sequence. We set $f_{i,e} = \lim_{n\to\infty} \psi_n^i(e)$ for all $e \in E$.

We claim that the numbers $(f_{i,e})_{i\in[k],e\in E}$ form a solution of (1)–(5). That (1) holds is obvious from the definition. (2) follows from the fact that $\sum_{e\in E} \lambda(n,e) = 1$ for all $n \in \mathbb{N}$. To show that (3) holds, fix $v \in V$. Note that we have $\sum_{e\in \mathrm{In}(v)} \kappa(n,e) - \sum_{e\in \mathrm{Out}(v)} \kappa(n,e) \in \{-1,0,1\}$ and therefore $-1/n \le \sum_{e\in \mathrm{In}(v)} \lambda(n,e) - \sum_{e\in \mathrm{Out}(v)} \lambda(n,e) \le 1/n$ for all $n \in \mathbb{N}$. Hence, the terms $\sum_{e\in \mathrm{In}(v)} \varphi_n^i(e) - \sum_{e\in \mathrm{Out}(v)} \varphi_n^i(e)$ converge to 0 when n goes to infinity. Since $\psi_0^i, \psi_1^i, \dots$ is a subsequence of $\varphi_0^i, \varphi_1^i, \dots$, the same is true for the terms $\sum_{e\in \mathrm{In}(v)} \psi_n^i(e) - \sum_{e\in \mathrm{Out}(v)} \psi_n^i(e)$. Since $\lim_{n\to\infty} \psi_n^i(e) = f_{i,e}$ exists for all $e \in E$, this implies that $\sum_{e\in \mathrm{In}(v)} f_{i,e} - \sum_{e\in \mathrm{Out}(v)} f_{i,e} = 0$, which proves (3). In order to prove (4) and (5), note that for all $i,j \in [k]$ we have

$$
\begin{aligned}
\phi_i(\pi) &= \liminf_{n\to\infty} r_i(\pi \restriction n)/n \\
&\le \liminf_{n\to\infty} r_i(\pi \restriction k_n^j)/k_n^j \\
&= \liminf_{n\to\infty} \sum_{e\in E} \lambda(k_n^j, e) \cdot r_i(e) \\
&= \liminf_{n\to\infty} \sum_{e\in E} \varphi_n^j(e) \cdot r_i(e) \\
&\le \lim_{n\to\infty} \sum_{e\in E} \psi_n^j(e) \cdot r_i(e) \\
&= \sum_{e\in E} f_{j,e} \cdot r_i(e)\,.
\end{aligned}
$$

Moreover, if $i = j$, both inequalities are equalities since $\lim_{n\to\infty} r_i(\pi \restriction k_n^i)/k_n^i$ exists and equals $\phi_i(\pi)$. Hence, $\sum_{e\in E} f_{i,e} \cdot r_i(e) = \phi_i(\pi) \le \sum_{e\in E} f_{j,e} \cdot r_i(e)$ for all $i,j \in [k]$, which proves (5). Finally, (4) follows from the assumption that $\bar{x} \le \phi(\pi) \le \bar{y}$. ☐

Lemma 28. For all $n \in \mathbb{N}$,

$$(n-2) \cdot \sum_{j=1}^{n-1} j! < n!$$

Proof. By induction over n. ☐

Lemma 29. Assume that G is strongly connected and that there exists a solution to (1)–(5). Then there exists an infinite path π in G such that $\bar{x} \le \phi(\pi) \le \bar{y}$.

Proof. Let G be strongly connected and assume that there exists a solution to (1)–(5). It is well-known that if a given system of linear constraints has a solution, then there exists one in rational numbers. Let $(f_{i,e})_{i\in[k],e\in E}$ be such a solution, where w.l.o.g. $f_{i,e} = c_{i,e}/d$ with $c_{i,e} \in \mathbb{N}$ and $d \in \mathbb{N} \setminus \{0\}$. Finally, let $\bar{z} \in \mathbb{R}^{[k]}$ be defined by $z_i = \sum_{e\in E} f_{i,e} \cdot r_i(e)$; by (4), $\bar{x} \le \bar{z} \le \bar{y}$. We claim that there exists an infinite path π in G with $\phi(\pi) = \bar{z}$.

For each $i \in [k]$ consider the directed multigraph G_i, which is derived from G by replacing a single edge $(u, v) \in E$ by as many as $c_{i,e}$ edges from u to v. By (3), we have $\sum_{e\in\mathrm{In}(v)} c_{i,e} = \sum_{e\in\mathrm{Out}(v)} c_{i,e}$ for all $v \in V$. Hence, in G_i each vertex has as many incoming edges as outgoing edges, which is a necessary and sufficient condition for the existence of an Eulerian cycle in each of the connected components of G_i. These cycles give rise to (disjoint, not necessarily simple) cycles $\gamma_1^i, \ldots, \gamma_m^i$ in G, where $m \le |V|$.

Consider for each $n \in \mathbb{N}$ the cycle ζ_n^i that starts by repeating the cycle γ_1^i n times, then takes the shortest path to the first vertex in the cycle γ_2^i, repeats this cycle n times, and so on, until, after repeating the cycle γ_m^i n times, taking the shortest path back to γ_1^i. Let $M = \max_{v\in V} r_j(v)$ be the maximum weight w.r.t. r_j. Note that:

$$
\begin{aligned}
&n \cdot \sum_{e\in E} c_{i,e} \cdot r_j(e) \le r_j(\zeta_n^i) \le n \cdot \sum_{e\in E} c_{i,e} \cdot r_j(e) + |V|^2 \cdot M,\\
&n \cdot \sum_{e\in E} c_{i,e} \le |\zeta_n^i| \le n \cdot \sum_{e\in E} c_{i,e} + |V|^2.
\end{aligned}
$$

Hence,

$$
\lim_{n\to\infty} \frac{r_j(\zeta_n^i)}{|\zeta_n^i|} = \frac{\sum_{e\in E} c_{i,e} \cdot r_j(e)}{\sum_{e\in E} c_{i,e}} = \frac{\sum_{e\in E} f_{i,e} \cdot r_j(e)}{\sum_{e\in E} f_{i,e}} = \sum_{e\in E} f_{i,e} \cdot r_j(e) \ge z_j,
$$

where the last inequality follows from (5). Moreover, if $i = j$, we have equality, i.e. $\lim_{n\to\infty} r_i(\zeta_n^i)/|\zeta_n^i| = z_i$.

The desired infinite path π is the concatenation of finite paths π_n, where $n = 1, 2 \ldots$. The path π_n repeats the cycle $\zeta_n^{n \bmod k}$ $n!$ times and then takes the shortest path to the first state on the cycle $\zeta_n^{(n+1) \bmod k}$. We will now prove that $\phi_0(\pi) = z_0$; for all other weight functions, the proof is analogous. For all $n \in \mathbb{N}$, we have:

$$
\begin{aligned}
&\sum_{j=1}^{nk} j!\, r_0\big(\zeta_j^{j \bmod k}\big) - nk|V|M \le r_0(\pi_1 \cdots \pi_{nk}) \le \sum_{j=1}^{nk} j!\, r_0\big(\zeta_j^{j \bmod k}\big) + nk|V|M,\\
&\sum_{j=1}^{nk} j!\, \big|\zeta_j^{j \bmod k}\big| \le |\pi_1 \cdots \pi_{nk}| \le \sum_{j=1}^{nk} j!\, \big|\zeta_j^{j \bmod k}\big| + nk|V|\,.
\end{aligned}
$$

By Lemma 28, we have $\lim_{n\to\infty} \sum_{j=1}^{nk-1} j!/(nk)! = 0$. Hence, and since $|\zeta_j^i|, r_0(\zeta_j^i) \leq j \cdot c$ for some constant c, we have:

$$\lim_{n\to\infty} \frac{1}{nk(nk)!} \cdot \sum_{j=1}^{nk} j!\, r_0\big(\zeta_j^{j \bmod k}\big) = \lim_{n\to\infty} \frac{1}{nk} \cdot r_0(\zeta_{nk}^0) = \sum_{j=1}^{m} r_0(\gamma_j^0),$$
$$\lim_{n\to\infty} \frac{1}{nk(nk)!} \cdot \sum_{j=1}^{nk} j!\, \big|\zeta_j^{j \bmod k}\big| = \lim_{n\to\infty} \frac{1}{nk} \cdot |\zeta_{nk}^0| = \sum_{j=1}^{m} |\gamma_j^0| \,.$$

Hence,

$$\lim_{n\to\infty} \frac{r_0(\pi_1 \cdots \pi_{nk})}{|\pi_1 \cdots \pi_{nk}|} = \lim_{n\to\infty} \frac{r_0(\zeta_{nk}^0)}{|\zeta_{nk}^0|} = z_0 \,.$$

We have thus found a subsequence of $r_0(\pi \restriction n)/n$ that converges to z_0, which implies that $\phi_0(\pi) = \liminf_{n\to\infty} r_0(\pi \restriction n)/n \leq z_0$. On the other hand, using the fact that $\lim_{n\to\infty} r_0(\zeta_n^i)/|\zeta_n^i| \geq z_0$ for all $i \in [k]$, we can show that $\phi_0(\pi) \geq z_0$. $\square$

Proof (of Theorem 18). Since the limit-average criterion is prefix-independent, it suffices to decompose G into its strongly connected components (which can be done in linear time) and check for each component C that is reachable from v_0 whether exists an infinite path in C with $\bar{x} \leq \phi(\pi) \leq \bar{y}$. By Lemmas 27 and 29, such a path exists if and only if there exists a solution to the linear constraints (1)–(5) derived from C. The existence of such a solution can be checked in polynomial time (see [23]). $\square$